

# YOUNG TOWERS FOR PRODUCT SYSTEMS

STEFANO LUZZATTO

*International Center for Theoretical Physics (ICTP)  
Strada Costiera 11, Trieste, Italy*

MARKS RUZIBOEV

*International School for Advanced Studies (SISSA)  
Via Bonomea 265, Trieste, Italy  
International Center for Theoretical Physics (ICTP)  
Strada Costiera 11, Trieste, Italy*

**ABSTRACT.** We show that the direct product of maps with Young towers admits a Young tower whose return times decay at a rate which is bounded above by the slowest of the rates of decay of the return times of the component maps. An application of this result, together with other results in the literature, yields various statistical properties for the direct product of various classes of systems, including Lorenz-like maps, multimodal maps, piecewise  $C^2$  interval maps with critical points and singularities, Hénon maps and partially hyperbolic systems.

## 1. INTRODUCTION AND STATEMENT OF RESULTS

**1.1. Product systems and Young towers.** Let  $f_i : M_i \rightarrow M_i$ ,  $i = 1, \dots, \ell$ , be a family of maps defined on a family of Riemannian manifolds. Define the product map  $f = f_1 \times \dots \times f_\ell$  on  $M = M_1 \times \dots \times M_\ell$  by

$$(1.1) \quad f(x_1, \dots, x_\ell) = (f_1(x_1), \dots, f_\ell(x_\ell)).$$

Dynamical properties of the product system  $f$  can be partially but not completely deduced from the dynamical properties of its components, as the product system may exhibit significantly richer dynamics. Even in the simplest setting of the product of two

---

*E-mail addresses:* luzzatto@ictp.it, mruziboe@sissa.it.

*Date:* 28 April 2015.

*2000 Mathematics Subject Classification.* 37A05, 37A25.

*Key words and phrases.* Decay of correlations, product systems, Young Towers.

We would like to thank Viviane Baladi and Carlangelo Liverani for several useful discussions and their help on some parts of the paper, and a referee for their comments which helped us significantly improve the presentation of the paper.

identical maps  $g \times g : M \rightarrow M$ , if  $p, q$  are periodic points for  $g$ , then the set-theoretic product of the periodic orbits  $\mathcal{O}^+(p) = \{p_1, \dots, p_m\}$  and  $\mathcal{O}^+(q) = \{q_1, \dots, q_n\}$  may consist of several periodic orbits<sup>1</sup>, which are in some sense *new* periodic orbits which do not exist in either of the original systems. Similarly, from an ergodic-theoretic point of view, the space  $\mathcal{M}_{g \times g}$  of invariant probability measures for the product map contains many invariant measures which are not products of invariant measures for  $g$  such as measures supported on the “new” periodic orbits mentioned above and measures supported on the “diagonal”  $\{(x, x) : x \in M\} \subset M \times M$  which is invariant for the product.<sup>2</sup>

One important tool that has emerged in the last decade or so for the study of dynamical and statistical properties of various kinds of dynamical systems, is that of *Young towers*, after the pioneering work of Young [49]. The direct product of Young towers is not in general itself a Young tower, and so it is not immediately obvious that the direct product of systems which admit a Young tower also admits a Young tower. The main result of this paper is to show that, nevertheless, a Young tower can be constructed for the product system (1.1) if each component  $f_i$  admits a Young tower, and that we can obtain some estimates for the decay of the return times of the tower for the product in terms of the rates of decay for the individual towers. We will also discuss various applications of this result to products of systems for which Young towers are known to exist.

**1.2. Statement of results.** We postpone the precise formal definition of a Young tower to Section 2 since it is somewhat technical and our main results can be stated without any reference to the details. The definition is slightly different depending on whether  $f$  is invertible or not but both cases we have a reference probability measure  $m$  on a “base”  $\Delta_0 \subset M$ , and a return time function  $R : \Delta_0 \rightarrow \mathbb{N}$ . We suppose that each of our maps  $f_i : M_i \rightarrow M_i$ ,  $i = 1, \dots, \ell$  admit Young towers with reference measures  $m_0^{(i)}$  defined on bases  $\Delta_0^{(i)}$  with return time functions  $R^{(i)} : \Delta_0^{(i)} \rightarrow \mathbb{N}$ . Let  $f : M \rightarrow M$  denote the corresponding product system (1.1), and  $\bar{m}_0 = m_0^{(1)} \times \dots \times m_0^{(\ell)}$  denote the product measure on  $\bar{\Delta}_0 = \Delta_0^{(1)} \times \dots \times \Delta_0^{(\ell)}$ . Finally, let

$$(1.2) \quad \mathcal{M}_n = \max_{i=1, \dots, \ell} m_0^{(i)} \{R^{(i)} > n\}.$$

We always assume that either all of the maps are invertible or all are noninvertible so that the product system is of the same type.

**Theorem.** *The product system  $f$  admits a tower with reference measure  $\bar{m}_0$  on  $\bar{\Delta}_0$  and return time  $T : \bar{\Delta}_0 \rightarrow \mathbb{N}$  satisfying the following bounds.*

**Exponential Decay:** *if  $\mathcal{M}_n = \mathcal{O}(e^{-\tau n})$  for some  $\tau > 0$ , then*

$$\bar{m}_0\{T > n\} = \mathcal{O}(e^{-\tau' n}).$$

<sup>1</sup> In fact, if  $\gcd\{m, n\} = 1$ , then  $(p, q)$  is periodic point of period  $mn$ . On the other hand, if  $\gcd\{m, n\} = k > 1$ , then  $(p, q)$  is periodic point with period  $mn/k$ , but  $\mathcal{O}^+(p) \times \mathcal{O}^+(q)$  is a union of  $k$  periodic orbits. As a simple example, consider the case  $m = 2$  and  $n = 4$ . In this case, the product of the orbits splits into two orbits for the product map:  $\mathcal{O}^+(p) \times \mathcal{O}^+(q) = \mathcal{O}^+(p_1 \times q_1) \cup \mathcal{O}^+(p_2 \times q_1)$ .

<sup>2</sup>Thanks to M. Blank for this interesting observation.

for some  $\tau' > 0$ .

**Stretched Exponential Decay** if  $\mathcal{M}_n = \mathcal{O}(e^{-\tau n^\theta})$  for some  $\tau, \theta > 0$ , then

$$\bar{m}_0\{T > n\} = \mathcal{O}(e^{-\tau' n^{\theta'}});$$

for all  $0 < \theta' < \theta$  and some  $\tau' = \tau'(\theta') > 0$ .

**Polynomial Decay** if  $\mathcal{M}_n = \mathcal{O}(n^{-\alpha})$  for some  $\alpha > \ell$ , then

$$\bar{m}_0\{T > n\} = \mathcal{O}(n^{\ell-\alpha}).$$

The idea of constructing a Young tower for a product of Young towers is sketched, in the non-invertible setting with exponential or polynomial return times, without a fully developed proof, in the PhD thesis of Vincent Lynch [34]. Our construction and estimates lead to a "loss" of one exponent for each component of the product system in the case of polynomial rates of decay, and also do not allow us to obtain results if the decay is slower than polynomial, such as in the interesting examples of Holland [25] which exhibit decay at rates of the form  $(\log \circ \dots \circ \log n)^{-1}$ . It is not clear to us if this is just a technical issue or if there might be some deeper reasons.

**1.3. Applications.** Young towers have been shown to imply a variety of statistical properties such as decay of correlations, invariance principles, limit theorems which in some cases can also be quantified in terms of the rate of decay of the tail of the return times associated to the tower [5, 18, 19, 23, 35, 36, 37, 49, 50]. An immediate consequence is that the dynamical systems which are direct products of systems which admit Young towers satisfy the statistical properties corresponding to the tail estimates of the product as given in our Theorem. Some examples of these systems include the following. *Lorenz-like interval maps* which are uniformly expanding and have a single singularity with dense preimages satisfying  $|f'(x)| \approx |x|^{-\beta}$  for some  $\beta \in (1/2, 1)$  admit a Young Tower with exponential tail, see [21] for the precise technical conditions; *Multimodal maps* for which the decay rate of the return times was obtained in terms of the growth rate of the derivative along the critical orbits [15]; *Maps with critical points and singularities* in one and also higher dimensions [3, 22, 7]; *Planar periodic Lorentz gas* was introduced by Sinai [44] and admits Young Towers with exponential tails [50, 19]; *Hénon maps*, for certain choices of parameters  $(a, b)$ , the maps  $H_{a,b} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  given by  $H_{a,b}(x, y) = (1 - ax^2 + y, bx)$  admit a Young Tower with exponential tail [11, 12]; *Partially hyperbolic systems* under certain additional conditions [6, 5].

We emphasize that the construction of Young towers, and especially the estimation of the decay of the return times, is in general highly non-trivial and relies on the specific geometric and dynamical properties of the system under consideration. The geometry of the direct product of any of the systems mentioned above is extremely complicated and it is doubtful that a tower construction could be achieved without taking advantage of the information that each component admits a tower with certain decay rates and applying our theorem.

In certain cases, combining our result with existing literature, it is possible to deduce statistical properties of the product system directly from the statistical properties of the component systems without assuming a priori that these admit Young towers. Indeed, in certain settings, such as that of non-uniformly expanding systems, statistical properties such as decay of correlations or large deviations imply the existence of Young tower with corresponding tail estimates see e.g. [2, Theorem 4.2]. Thus, taking the direct product of any finite number of such systems we can apply our result and those of [36] to conclude that Large Deviations for the component systems implies Large Deviations for the product system. The same argument also works for Decay of Correlations but in this case, as Carlangelo Liverani pointed out to us, it is possible to give an elementary and self-contained direct proof. Since it does not appear to exist in the literature, we include it in the Appendix of this paper.

**1.4. Overview of the paper.** In Section 2 we give the precise formal definition of a Young tower in the non-invertible setting. Most of the paper will then be dedicated to proving our main result in this setting. In Section 7 we will give the formal definition in the invertible case and show how the construction easily extends to this setting. In Section 3 we give the full combinatorial/topological construction of the tower for the product system. In Section 4 we give some basic estimates concerning the asymptotics of the tail of the newly constructed tower for the product, and in Section 5 we apply these to get specific estimates in the exponential, stretched exponential and polynomial case. In Section 6 combining the results from previous sections, we finalize proof of the Theorem in the non-invertible case. Finally, in the appendix we prove some technical estimates and explain the estimate due to Liverani on the decay of correlations for product systems.

## 2. YOUNG TOWERS

**2.1. Gibbs-Markov-Young Towers.** We start with the formal definition of Young Tower for non-invertible maps. To distinguish this case from the tower for invertible maps, this structure sometimes is referred to as a Gibbs-Markov-Young (GMY) structure or GMY-tower. Let  $f : M \rightarrow M$  be a  $C^{1+\alpha}$ ,  $\alpha \in (0, 1)$  local diffeomorphism (outside some critical/singular set) of a Riemannian manifold  $M$  on which we have a Riemannian volume which we will refer to as Lebesgue measure. To define a GMY-structure for  $f$  we start with a measurable set  $\Delta_0$  with finite positive Lebesgue measure, we let  $m_0$  denote the restriction of Lebesgue measure to  $\Delta_0$ , a mod 0 partition  $\mathcal{P} = \{\Delta_{0,i}\}$  of  $\Delta_0$ , and a return time function  $R : \Delta_0 \rightarrow \mathbb{N}$  that is constant on the partition elements, i.e.  $R|_{\Delta_{0,i}} = R_i$  such that the induced map  $f^R : \Delta_0 \rightarrow \Delta_0$  is well defined. Then, for any two point  $x, y \in \Delta_0$  we can define the *separation time*  $s(x, y)$  as the smallest  $k \geq 0$  such that  $(f^R)^k(x)$  and  $(f^R)^k(y)$  lie in different partition elements and assume that there exists  $\beta \in (0, 1)$  and  $D > 0$  such that the following conditions are satisfied.

(G1) **Markov:** for any  $\Delta_{0,i} \in \mathcal{P}$  the map  $f^{R_i} : \Delta_{0,i} \rightarrow \Delta_0$  is a bijection.

(G2) **Uniform expansion:**  $\|(Df^R)^{-1}(x)\| \leq \beta$  for  $m_0$  a.e.  $x$ .

(G3) **Bounded distortion:** for a.e. pair  $x, y \in \Delta_0$  with  $s(x, y) < \infty$  we have

$$(2.1) \quad \left| \frac{\det Df^R(x)}{\det Df^R(y)} - 1 \right| \leq D\beta^{s(f^R x, f^R y)}.$$

(G4) **Integrability:**  $\int R dm_0 < \infty$ .

(G5) **Aperiodicity:**  $\gcd\{R_i\} = 1$ .

Young showed that the first three assumptions (G1)-(G3) imply that the map  $f$  admits an  $f$ -invariant measure  $\mu$  which is absolutely continuous with respect to Lebesgue. Condition (G4) implies that this measure is finite, and therefore can be taken to be a probability measure, and condition (G5) implies that it is mixing. We note that the integrability and the aperiodicity assumption are not always included in the definition of a Young Tower. We include them here because we actually require the existence of a mixing probability measure for each component for our argument to work in the proof of the Theorem and also because we are interested in applications to the problem of the rate of decay of correlations for product systems and this requires the measures involved to be mixing probability measures.

We note also that the return time  $R$  is not generally a *first return* time of point to  $\Delta_0$ . It is therefore often useful to work with an “extension” of  $f$  in which the returns to  $\Delta_0$  are first return times. This extension is precisely what we refer to as a *GMV-tower*. The formal construction of this extension proceeds as follows. We let

$$(2.2) \quad \Delta = \{(z, n) \in \Delta_0 \times \mathbb{Z}_0^+ \mid R(z) > n\},$$

where  $\mathbb{Z}_0^+$  denotes the set of all nonnegative integers. For  $\ell \in \mathbb{Z}_0^+$  the subset  $\Delta_\ell = \{(\cdot, \ell) \in \Delta\}$  of  $\Delta$  is called its  $\ell$ th level. By some slight abuse of notation, we let  $\Delta_0$  denote both the subset of the Riemannian manifold  $M$  on which the induced map  $f^R$  is defined and the 0th level of tower  $\Delta$ . The collection  $\Delta_{\ell,i} := \{(z, \ell) \in \Delta_\ell \mid (z, 0) \in \Delta_{0,i}\}$  forms a partition of  $\Delta$  that we denote by  $\eta$ . The set  $\Delta_{R_i-1,i}$  is called the *top level* above  $\Delta_{0,i}$ . We can then define a map  $F : \Delta \rightarrow \Delta$  letting

$$(2.3) \quad F(z, \ell) = \begin{cases} (z, \ell + 1) & \text{if } \ell + 1 < R(z), \\ (f^{R(z)}(z), 0) & \text{if } \ell + 1 = R(z). \end{cases}$$

There exists a natural projection  $\pi : \Delta \rightarrow M$  defined by  $\pi(x, \ell) = f^\ell(x_0)$  for  $x \in \Delta$  with  $F^{-\ell}(x) = x_0 \in \Delta_0$ . Notice that  $\pi$  is a semi-conjugacy  $f \circ \pi = \pi \circ F$ .

For future reference we note that we can extend the return time and the separation time to all of  $\Delta$ . Indeed, for any  $x \in \Delta$  we can define a *first hitting time* by

$$(2.4) \quad \hat{R}(x) := \min\{n \geq 0 : F^n(x) \in \Delta_0\}.$$

Notice that if  $x \in \Delta_0$  then  $\hat{R}(x) = 0$ . We also extend the separation time to  $\Delta$  by setting  $s(x, y) = 0$  if  $x$  and  $y$  belong to different elements of  $\eta$  and  $s(x, y) = s(\tilde{x}, \tilde{y})$  if  $x$  and  $y$  are in the same element of  $\eta$ , where  $\tilde{x}$  and  $\tilde{y}$  are the corresponding projections to  $\Delta_0$ .

We define a reference measure  $m$  on  $\Delta$  as follows. Let  $\mathcal{A}$  be the Borel  $\sigma$ -algebra on  $\Delta_0$  and let  $m_0$  denote the restriction of Lebesgue measure to  $\Delta_0$  where, as mentioned

above, we are identifying  $\Delta_0 \subset M$  with the 0'th level of the tower. For any  $\ell \geq 0$  and  $A \subset \Delta_\ell$  such that  $F^{-\ell}(A) \in \mathcal{A}$  define  $m(A) = m_0(F^{-\ell}(A))$ . Notice that with this definition, the restriction of  $m$  to  $\Delta_0$  is exactly  $m_0$ , whereas the restriction of  $m$  to the upper levels of the tower is not equal to the Riemannian volume of their projections on  $M$  because the tower map  $F : \Delta \rightarrow \Delta$  is by definition an isometry between one level and the next, except on the top level where it maps back to the base with an expansion which corresponds to the “accumulated” expansion of all the iterates on the manifold. Correspondingly, for every  $x \in \Delta_{R_i-1,i} \subset \Delta$ , the Jacobian of  $F$  at  $x$  is  $JF(x) = \det Df^R(x_0)$ , where  $x_0 = F^{-R_i+1}x \in \Delta_0$ , and  $JF(x) = 1$  otherwise.

### 3. A TOWER FOR THE PRODUCT

In this section we begin the proof of our main result. To simplify the notation we will assume that we have a product of only two systems, the general case follows immediately by iterating the argument.

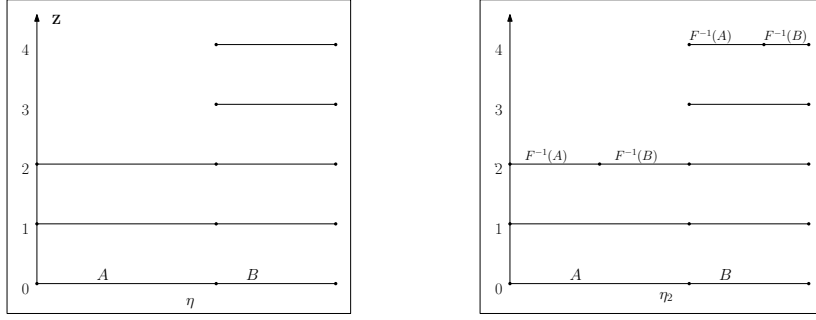
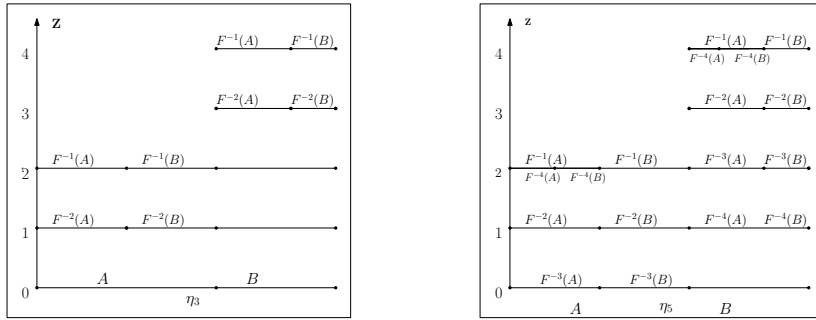
**3.1. Basic ideas and notation.** We begin by introducing some basic notions. Suppose  $F : (\Delta, m) \circlearrowleft$  is a GMY-tower as defined in (2.2) and (2.3) above and let  $\eta$  be the partition of  $\Delta$  into  $\Delta_{\ell,i}$ 's. Then, for  $n \geq 1$ , let

$$(3.1) \quad \eta_n := \bigvee_{j=0}^{n-1} F^{-j}\eta := \{A_1 \cap F^{-1}(A_2) \cap \dots \cap F^{1-n}(A_n) \mid A_1, \dots, A_n \in \eta\}$$

be the refinements of the partition  $\eta$  defined by the map  $F$ . For  $x \in \Delta$  let  $\eta_n(x)$  be the element containing  $x$ . From (3.1) it is easily seen that  $\eta_n(x)$  has the form

$$\eta_n(x) = \left( \bigvee_{j=0}^{n-1} F^{-j}\eta \right) (x) = \eta(x) \cap F^{-1}\eta(F(x)) \cap \dots \cap F^{1-n}\eta(F^{n-1}(x)).$$

*Remark 1.* To get a better feeling for the partitions  $\eta_n$  notice that from the definition of tower for  $x \notin \Delta_0$  we have  $F^{-1}(\eta(x)) = \eta(F^{-1}(x))$ , which shows that the element  $\eta(x)$  gets refined only when  $F^{-j}(x) \in \Delta_0$  for some  $j$ ,  $j = 0, \dots, n-2$ . It may be instructive to consider more in detail the cases  $n = 2$  and  $n = 3$  (notice that  $\eta_1 = \eta$ ). For  $n = 2$ , from (3.1) we have  $F^{-1}\eta \vee \eta$ . In this case only the elements on the top levels (recall definition just before equation (2.3)) get refined so that the new elements are mapped by  $F$  bijectively onto  $\Delta_{0,i} \subset \Delta_0$ , for some  $i$ . All the other elements remain unchanged, see Figure 1 (for simplicity, the pictures are drawn when the partition of base contains only two elements and has return times 3 and 5, in particular, for  $x \in \Delta$  with  $F^2(x) \in \Delta_0$  we have  $F^2(\eta_2(x)) = \Delta_0$ . This is because  $\eta_2(x) = \eta(x)$  and  $F(\eta(x))$  is top level element of  $\eta$ ). For  $n = 3$  all the elements on the top levels and on the levels “immediately below” the top levels get refined so that the top level elements of the new partition are mapped onto some  $\Delta_{0,i}$  by  $F$ ; the elements belonging to the levels immediately below the top levels are mapped onto some  $\Delta_{0,i}$  by  $F^2$  and other elements remain unchanged, see the left hand side of Figure 2. On the right hand side of Figure 2 is illustrated the

FIGURE 1.  $\eta$  and  $\eta_2$ FIGURE 2.  $\eta_3$  and  $\eta_5$ 

situation for  $n = 5$  in this simple example, where  $n = \min_i \{R_i\} + 2$ , and therefore the refinement procedure “reaches”  $\Delta_0$ . After this time, the top levels undergo a “second round” of refinements. In Figure 2 the bold elements of  $\eta_5$  are the elements of  $\eta$  which have been refined twice. In general, for each  $n$  the refinement procedure affect  $n - 1$  levels below the top levels. If  $n$  is sufficiently large, some of the partition elements might get refined several times.

The following statement follows almost immediately from the observation above. Recall also the definition of the first hitting time in (2.4).

**Lemma 3.1.** For any  $x \in F^{-n}(\Delta_0)$  the map  $F^n : \eta_n(x) \rightarrow \Delta_0$  is a bijection and  $F^n(\eta(x)) = \Delta_0$ . Moreover if  $\hat{R}(x) = n$  then  $\hat{R}|_{\eta_n(x)} \equiv n$ .

*Proof.* The proof is by induction on  $n$ . For  $n = 1$  we have no refinement,  $\eta_1 = \eta$ , and therefore the conclusion follows from the definition of tower. Assume that the assertion is true for  $n = k$ . From equality (3.1) we obtain  $\eta_{k+1} = F^{-1}(\eta_k) \vee \eta$ . Let  $x \in \Delta$  be a point, such that  $F^{k+1}(x) \in \Delta_0$  then using the relation  $\eta_{k+1}(x) = F^{-1}\eta_k(F(x)) \cap \eta(x)$  we obtain  $F^{k+1}(\eta_{k+1}(x)) = F^k(\eta_k(F(x))) \cap F^{k+1}(\eta(x))$ . Since  $F^k(\eta_k(F(x))) = \Delta_0$  from the inductive assumption and  $\Delta_0 \subset F^{k+1}(\eta(x))$  from the definition of tower we

get  $F^{k+1}(\eta_{k+1}(x)) = \Delta_0$ . Since the first return time is constant on the elements of  $\eta$ , the second assertion follows.  $\square$

**3.2. Return times to  $\bar{\Delta}_0$ .** We are now ready to begin the construction of the tower for the product system. Since we are considering the product of just two systems, we will omit superfluous indexing and let  $f$  and  $f'$  be two maps that admit GMY-structure with the bases  $\Delta_0, \Delta'_0$  and return time functions  $R, R'$  respectively. Then we have associated GMY-towers

$$F : (\Delta, m) \curvearrowright \quad \text{and} \quad F' : (\Delta', m') \curvearrowright$$

with bases  $\Delta_0$  and  $\Delta'_0$  and return time functions  $R(x)$  and  $R'(x')$  respectively. Let

$$\bar{\Delta} = \Delta \times \Delta' \quad \text{and} \quad \bar{\Delta}_0 = \Delta_0 \times \Delta'_0$$

denote the product of the two towers, and the product of their bases respectively. Letting  $m_0, m'_0$  denote the restrictions of  $m, m'$  to  $\Delta_0, \Delta'_0$  respectively, we let  $\bar{m} = m \times m'$  and  $\bar{m}_0 = m_0 \times m'_0$  denote the product measures on the corresponding products. The direct product map  $\bar{F} = F \times F'$  is defined on  $\bar{\Delta}$  and we will construct a tower for  $\bar{F}$  with base  $\bar{\Delta}_0$ . We start by defining the return time function  $T$  on  $\bar{\Delta}_0$ . From Theorem 1 in [50] there exist mixing invariant probability measures  $\mu, \mu'$  for  $F$  and  $F'$ , equivalent to  $m, m'$  respectively, and with densities which are uniformly bounded above and below. Therefore, there exist constants  $c > 0$  and  $n_0 > 0$  such that

$$(3.2) \quad m(F^{-n}(\Delta_0) \cap \Delta_0) > c > 0 \quad \text{and} \quad m'(F'^{-n}(\Delta'_0) \cap \Delta'_0) > c > 0$$

for all  $n \geq n_0$ . We choose such  $n_0$  and introduce a sequence  $\{\tau_i\}$  of positive integers as follows. For  $\bar{x} = (x, x') \in \bar{\Delta}$  let

$$\tau_0(\bar{x}) = 0 \quad \text{and} \quad \tau_1(\bar{x}) := n_0 + \hat{R}(F^{n_0}x).$$

The other elements of the sequence  $\{\tau_i\}$  are defined inductively by iterating  $F$  or  $F'$  alternately depending on whether  $i$  is odd or even. More formally, for every  $j \geq 1$  let

$$(3.3) \quad \begin{aligned} \tau_{2j}(\bar{x}) &:= \tau_{2j-1}(\bar{x}) + n_0 + \hat{R}'(F'^{n_0+\tau_{2j-1}}x'), \\ \tau_{2j+1}(\bar{x}) &:= \tau_{2j}(\bar{x}) + n_0 + \hat{R}(F^{n_0+\tau_{2j}}x). \end{aligned}$$

*Remark 2.* Notice that at every step we “wait” for  $n_0$  iterates before defining the next term of the sequence. This implies that for any  $i$  we have  $\tau_i - \tau_{i-1} \geq n_0$  and therefore, from (3.2), we get

$$m_0(\Delta_0 \cap F^{\tau_{i-1}-\tau_i}\Delta_0) \geq c > 0 \quad \text{and} \quad m'_0(\Delta'_0 \cap F'^{\tau_{i-1}-\tau_i}\Delta'_0) \geq c > 0$$

We use this fact in the proof of the first item in Proposition 4.1.



**3.3. Initial step of construction of the partition of  $\bar{\Delta}_0$ .** Now we can begin to define a partition  $\bar{\eta}$  of  $\bar{\Delta}_0$ . Recall that the towers  $\Delta, \Delta'$  of the component systems admit by definition partitions into sets of the form  $\Delta_{\ell,j}, \Delta'_{\ell,j}$ . We will denote these given partitions by  $\eta, \eta'$  respectively and their restrictions to  $\Delta_0, \Delta'_0$  by  $\eta_0, \eta'_0$ . We let

$$\xi_0 = \eta_0 \times \eta'_0$$

denote the corresponding partition of the product  $\bar{\Delta}_0$ . Our goal is to define a partition  $\bar{\eta}$  of  $\bar{\Delta}_0$  with the property that for each  $\bar{x} \in \bar{\Delta}_0$  the corresponding partition element  $\bar{\eta}(\bar{x}) \in \bar{\eta}$  maps bijectively to  $\bar{\Delta}_0$  with some return time  $T(\bar{\eta}(\bar{x}))$ . Its construction requires the definition of an increasing sequence of partitions of  $\bar{\Delta}_0$  denoted by

$$\xi_0 \prec \xi_1 \prec \xi_2 \prec \xi_3 \prec \dots$$

A key property of these partitions will be that the sequences  $\tau_1, \dots, \tau_i$  are constant on elements of  $\xi_i$ . Moreover, the construction of  $\bar{\eta}$  implies that all return times  $T(\bar{x}) = T(\bar{\eta}(\bar{x}))$  are of the form  $T(\bar{x}) = \tau_i(\bar{x})$  for some  $i$  where  $\tau_i$  belongs to the sequence defined above. This property will be used below for the estimates of the tail of the return times. In particular, notice also that there may be some elements of  $\eta_0 \times \eta'_0$  which map bijectively to  $\bar{\Delta}_0$ , and thus are candidates for elements of  $\bar{\eta}$ , but are not guaranteed to satisfy the requirement just stated above. The partition  $\bar{\eta}$  will be defined as the union

$$(3.4) \quad \bar{\eta} = \bigcup_{i=1}^{\infty} \bar{\eta}_i$$

of disjoint sets  $\bar{\eta}_i$  which consists of a collection of subsets which are defined in the first  $i$  steps of the construction.

The systematic construction proceeds as follows. For each  $\bar{x} = (x, x') \in \bar{\Delta}_0$ , let

$$(3.5) \quad \eta_{\tau_1(x)}(x) := \left( \bigvee_{j=0}^{\tau_1(x)-1} F^{-j}\eta \right)(x).$$

It follows immediately that  $\eta_{\tau_1(x)}(x) \subseteq \eta_0(x)$ . Indeed, as the following simple Lemma proves, collection  $\eta_{\tau_1} := \{\eta_{\tau_1(x)}(x) | x \in \Delta_0\}$  is in fact a partition of  $\Delta_0$ .

**Lemma 3.2.**  $\eta_{\tau_1}$  is a partition of  $\Delta_0$ ,  $\eta_0 \prec \eta_{\tau_1}$  and  $\tau_1$  is constant on elements of  $\eta_{\tau_1}$ .

*Proof.* First of all, note that since  $\tau_1$  depends only on the first coordinate  $\eta_{\tau_1(x)}(x) \in \eta_n$ , with  $n = \tau_1(x)$ . Then Lemma 3.1 implies that  $F^{\tau_1(x)}(\eta_{\tau_1(x)}) = \Delta_0$  bijectively. In particular, all the points in  $\eta_{\tau_1(x)}(x)$  have the same combinatorics up to time  $\tau_1(x)$ , and hence  $\tau_1$  is constant on  $\eta_{\tau_1(x)}(x)$ .

Now, to prove the Lemma, since  $\eta_{\tau_1}$  clearly covers  $\Delta_0$ , we just need to show every pair of sets in  $\eta_{\tau_1}$  are either disjoint or coincide. Let  $\eta_{\tau_1(x)}(x)$  and  $\eta_{\tau_1(y)}(y)$  be two arbitrary elements of  $\eta_{\tau_1}$ . If  $\tau_1(x) = \tau_1(y) = n$  then  $\eta_{\tau_1(x)}(x)$  and  $\eta_{\tau_1(y)}(y)$  are the elements of  $\eta_n$ , hence they are disjoint or coincide. If  $\tau_1(x) \neq \tau_1(y)$  then  $\eta_{\tau_1(x)}(x) \cap \eta_{\tau_1(y)}(y) = \emptyset$  because  $\tau_1$  is constant on elements of  $\eta_{\tau_1}$ .  $\square$

We can now define the partition  $\xi_1$  of  $\bar{\Delta}_0$  by letting, for every  $\bar{x} = (x, x') \in \bar{\Delta}_0$ ,

$$\xi_1(\bar{x}) := \eta_{\tau_1(x)}(x) \times \eta'_0(x').$$

Notice that each element  $\Gamma \in \xi_1$  has an associated value of  $\tau_1$  such that  $F^{\tau_1}(x) \in \Delta_0$  for every  $x \in \pi\Gamma$ . On the other hand, we do not a priori have any information about the location of  $F^{\tau_1}(x')$  for  $x' \in \pi'\Gamma$  (since  $\tau_1$  is defined in terms of properties of  $F$ ). To study the distribution of such images, for a given  $\Gamma \in \xi_1$ , we consider sets of the form

$$(3.6) \quad \eta'_{\tau_1}(x') = \left( \bigvee_{j=0}^{\tau_1-1} F'^{-j}\eta' \right)(x').$$

The collection of such sets, for all  $x' \in \pi'\Gamma$ , form a refinement of  $\pi'\Gamma$ . For those points  $x' \in \pi'\Gamma$  such that  $F^{\tau_1}(x') \in \Delta'_0$  we then have that

$$F'^{\tau_1} : \eta'_{\tau_1}(x') \rightarrow \Delta'_0$$

bijectively. In this case we consider the set

$$\bar{\eta}_1(x, x') = \eta_{\tau_1}(x) \times \eta'_{\tau_1}(x').$$

which maps bijectively to  $\bar{\Delta}_0$  by  $\bar{F}^{\tau_1}$ , and let  $\bar{\eta}_1$  denote the collection of all sets of this form constructed at this step. This is the first collection of sets which will be included in the union (3.4) defined above.

**3.4. General step.** We now describe the general inductive step in the construction of the sequence of partitions  $\xi_i$  and the sets  $\bar{\eta}_i$ . The main inductive assumption is that partitions  $\xi_i$  of  $\bar{\Delta}_0$  have been constructed for all  $i < k$  in such a way that on each element of  $\xi_i$  the functions  $\tau_1, \dots, \tau_i$  are constant and such that each component of  $\bar{\eta}_{i-1}$  is contained inside an element of  $\xi_{i-1}$  and contains one or more elements of  $\xi_i$  (in particular elements of the partition  $\xi_i$  either have empty intersection with  $\bar{\eta}_{i-1}$  or are fully contained in some component of  $\bar{\eta}_{i-1}$ ).

The notation is slightly different depending on whether  $k$  is even or odd, according to the different definitions of  $\tau_k$  in these two cases, recall (3.3). For definiteness we assume that  $k$  is odd, the construction for  $k$  even is the same apart from the change in the role of the first and second components. We fix some  $\Gamma \in \xi_{k-1}$  and define the partition  $\xi_k|_{\Gamma}$  as follows. For  $\bar{x} \in \Gamma$ , let

$$\eta_{\tau_k(x)}(\bar{x}) = \left( \bigvee_{j=0}^{\tau_k(\bar{x})-1} F^{-j}\eta \right)(x).$$

A direct generalization of Lemma 3.2 gives that the collection of sets  $\eta_{\tau_k} := \{\eta_{\tau_k(x)}(x) | x \in \pi\Gamma\}$  is a partition of  $\pi\Gamma$  on whose elements  $\tau_k$  is constant. For every  $\bar{x} \in \Gamma$  we let

$$\xi_k(\bar{x}) := \eta_{\tau_k(x)}(\bar{x}) \times \pi'\Gamma.$$

This completes the definition of the partition  $\xi_k$  and allows us define  $\bar{\eta}_k$ . As mentioned in the inductive assumptions above, each component of  $\bar{\eta}_k$  will be contained in an element of  $\xi_k$ . Thus, generalizing the construction of such elements in the first step given above,

we fix one element  $\Gamma \in \xi_k$  and proceed as follows. By construction we have  $F^{\tau_k}(\pi\Gamma) = \Delta_0$  and  $F'^{\tau_k}(\pi'\Gamma)$  is spread around  $\Delta'$ . Therefore, for every  $x' \in \pi'\Gamma$  we consider sets of form

$$(3.7) \quad \eta'_{\tau_k}(x') = \left( \bigvee_{j=0}^{\tau_k-1} F'^{-j}\eta' \right)(x').$$

The collection of such sets, for all  $x' \in \pi'\Gamma$ , form a refinement of  $\pi'\Gamma$ . For those points  $x' \in \pi'\Gamma$  such that  $F^{\tau_k}(x') \in \Delta'_0$  we then have that

$$F'^{\tau_k} : \eta'_{\tau_k}(x') \rightarrow \Delta'_0$$

bijectively. In this case we consider the set

$$\bar{\eta}_k(x, x') = \eta_{\tau_k}(x) \times \eta'_{\tau_k}(x')$$

which maps bijectively to  $\bar{\Delta}_0$  by  $\bar{F}^{\tau_k}$ , and let  $\bar{\eta}_k$  denote the collection of all sets of this form constructed at this step. Moreover, for each such set we let  $T = \tau_k$ . Finally, notice that the construction of  $\bar{\eta}_k$  through the formula (3.7) implies that the elements of the partition  $\xi_{k+1}$ , to be constructed in the next step, are either disjoint from  $\bar{\eta}_k$  or contained components of  $\bar{\eta}_k$ . Indeed, the construction of  $\xi_{k+1}$  involves a formula analogous to (3.7) with  $\tau_k$  replaced by  $\tau_{k+1}$ , clearly yielding a finer partition of  $\pi'\Gamma$ .

This completes the general step of the construction and in particular allows us to define the set  $\bar{\eta}$  as in (3.4). We remark that by construction all components of  $\bar{\eta}$  are pairwise disjoint but we have not yet proved that  $\bar{\eta}$  is a partition of  $\bar{\Delta}_0$ . This will be an immediate consequence of Proposition 4.1 in the next section. For formal consistency of notation we let  $T = \infty$  on all points in the complements of the elements of  $\bar{\eta}$ . Notice that it follows from the construction that for any  $\Gamma \in \bar{\eta}$ :

- (1)  $T|\Gamma = \tau_i|\Gamma$  for some  $i$ .
- (2)  $\bar{F}^T(\Gamma) = \bar{\Delta}_0$ .

In particular all elements of  $\bar{\eta}$  satisfy properties (G1) and (G2) in the definition of Young Tower. In Section 4 we obtain some preliminary estimates concerning the general asymptotics of the return times of the elements of  $\bar{\eta}$  defined above, and in Section 5 we consider the specific cases of polynomial, stretched exponential and exponential decay rates. In Section 6 we use all these estimates to prove that  $\bar{\eta}$  is indeed a partition of  $\bar{\Delta}_0$  and that the return map  $\bar{F}^T$  satisfies the required properties (G3)-(G5).

#### 4. RETURN TIME ASYMPTOTICS

In this section we begin the study of the asymptotics of the return time  $T$ . We will use the following general notation for conditional measures: if  $\mu$  is a measure we write  $\mu(B|A) := \mu(A \cap B)/\mu(A)$ . Also, for every  $n \geq 1$ , we write

$$(4.1) \quad \bar{\mathcal{M}}_n := \sum_{j \geq n} \mathcal{M}_j$$

where  $\mathcal{M}_j$  is the bound on the tails of the component systems, as in (1.2). The main result of this section is the following

**Proposition 4.1.** *There exist constants  $\varepsilon_0, K_0 > 0$  such that for any  $i \geq 2$*

- (1)  $\bar{m}_0\{T = \tau_i | T > \tau_{i-1}\} \geq \varepsilon_0$
- (2)  $\bar{m}_0\{\tau_{i+1} - \tau_i \geq n | \Gamma\} \leq K_0 \bar{\mathcal{M}}_{n-n_0}$ , for any  $n > n_0$  and  $\Gamma \in \xi_i$ .

Recall that we have set  $T = \infty$  on the complement of points belonging to some element of  $\bar{\eta}$ , and notice that  $\{T > \tau_{i-1}\}$  is the set of points in  $\bar{\Delta}_0$  which do not belong to any elements of  $\eta_j$  for any  $j = 1, \dots, i-1$  (by some slight abuse of notation we could write this as  $\{T > \tau_{i-1}\} = \bar{\Delta}_0 \setminus \bigcup_{j=1}^{i-1} \bar{\eta}_j$ ). As an immediate consequence of item (1) we get the statement that  $T$  is finite for almost every point in  $\bar{\Delta}_0$  and therefore the collection of sets  $\bar{\eta}$  as in (3.4) is indeed a partition of  $\bar{\Delta}_0 \bmod 0$ .

The proof of Proposition 4.1 relies on some standard combinatorial estimates which, for completeness and to avoid interrupting the flow of the calculations, we include in Appendix A.

*Proof of (1).* By construction,  $\{T > \tau_{i-1}\}$  is a union of elements of the partition  $\xi_i$ . Thus

$$\bar{m}_0\{T = \tau_i | T > \tau_{i-1}\} = \frac{1}{\bar{m}_0\{T > \tau_{i-1}\}} \sum_{\Gamma \in \xi_i, T|_{\Gamma} > \tau_{i-1}} \bar{m}_0\{\{T = \tau_i\} \cap \Gamma\}.$$

Thus it is sufficient to prove  $\bar{m}_0\{T = \tau_i | \Gamma\} \geq \varepsilon_0$  for any  $\Gamma \in \xi_i$  on which  $T|_{\Gamma} > \tau_{i-1}$ . Assume for a moment  $i$  is even and let  $\Omega = \pi(\Gamma)$ ,  $\Omega' = \pi'(\Gamma)$ . Then by construction  $\pi'(\{T = \tau_i\} \cap \Gamma) = \Omega'$  and

$$\bar{m}_0\{T = \tau_i | \Gamma\} = \frac{\bar{m}_0(\{T = \tau_i\} \cap \Gamma)}{\bar{m}_0(\Gamma)} = \frac{m_0(\Omega \cap F^{-\tau_i} \Delta_0) m'_0(\Omega')}{m_0(\Omega) m'_0(\Omega')} = \frac{m_0(\Omega \cap F^{-\tau_i} \Delta_0)}{m_0(\Omega)}.$$

Now recall that  $F^{\tau_{i-1}}(\Omega) = \Delta_0$ , which implies  $\Omega \cap F^{-\tau_{i-1}} \Delta_0 = \Omega$  and therefore

$$\frac{m_0(\Omega \cap F^{-\tau_i} \Delta_0)}{m_0(\Omega)} = \frac{m_0(\Omega \cap F^{-\tau_{i-1}}(\Delta_0 \cap F^{\tau_{i-1}-\tau_i} \Delta_0))}{m_0(\Omega)} = F_*^{\tau_{i-1}}(m_0|_{\Omega})(\Delta_0 \cap F^{\tau_{i-1}-\tau_i} \Delta_0).$$

Notice that  $m_0(\Delta_0 \cap F^{\tau_{i-1}-\tau_i} \Delta_0) \geq c > 0$  since  $\tau_i - \tau_{i-1} \geq n_0$ . Letting  $\nu = F_*^{\tau_{i-1}}(m_0|_{\Omega})$ , applying Corollary A.5 with  $n = \tau_{i-1}$  for  $x, y \in \Delta_0$  we get

$$\bar{m}_0\{T = \tau_i | \Gamma\} = \nu(\Delta_0 \cap F^{\tau_{i-1}-\tau_i} \Delta_0) \geq \frac{m_0(\Delta_0 \cap F^{\tau_{i-1}-\tau_i} \Delta_0)}{(1+D)m_0(\Delta_0)} \geq \frac{c}{(1+D)m_0(\Delta_0)}.$$

For odd  $i$ 's we can just change  $F$  to  $F'$  and do all the calculations, that gives the estimate

$$\bar{m}_0\{T = \tau_i | \Gamma\} > \frac{c}{(1+D')m'_0(\Delta'_0)}.$$

Taking  $\varepsilon_0 = c \min\{\frac{1}{(1+D)m_0(\Delta_0)}, \frac{1}{(1+D')m'_0(\Delta'_0)}\}$  we get the assertion.  $\square$

*Proof of (2).* Assume for a moment  $i$  is even and let, as above,  $\Omega = \pi(\Gamma)$ ,  $\Omega' = \pi'(\Gamma)$ . Since  $\tau_i$  is constant on the elements of  $\xi_i$  we have

$$\pi \left( \{ \bar{x} = (x, x') | \hat{R} \circ F^{\tau_i+n_0}(x) > n \} \cap \Gamma \right) = \{ x | \hat{R} \circ F^{\tau_i+n_0}(x) > n \} \cap \Omega.$$

For convenience we begin by estimating  $\bar{m}_0\{\tau_{i+1} - \tau_i - n_0 > n | \Gamma\}$ . From the definition of  $\tau_i$ , letting  $\nu = F_*^{\tau_i-1}(m_0 | \Omega)$ , we have

$$\begin{aligned} \bar{m}_0\{\tau_{i+1} - \tau_i - n_0 > n | \Gamma\} &= \bar{m}_0\{\hat{R} \circ F^{\tau_i+n_0} > n | \Gamma\} = \frac{\bar{m}(\{\hat{R} \circ F^{\tau_i+n_0} > n\} \cap \Gamma)}{\bar{m}(\Gamma)} \\ &= \frac{m'_0(\Omega') m_0(\pi\{\hat{R} \circ F^{\tau_i+n_0} > n\} \cap \Omega)}{m_0(\Omega) m'_0(\Omega')} \\ &= \frac{m_0(\{\hat{R} \circ F^{\tau_i+n_0} > n\} \cap \Omega)}{m_0(\Omega)} = m_0\{\hat{R} \circ F^{\tau_i+n_0} > n | \Omega\} \\ &= F_*^{\tau_i+n_0} m_0\{\hat{R} > n | \Omega\} = F_*^{\tau_i-\tau_{i-1}+n_0} \nu\{\hat{R} > n\}. \end{aligned}$$

To bound the final term in terms of  $m_0\{\hat{R} > n\}$  it is sufficient to show that the density of  $F_*^{\tau_i-\tau_{i-1}+n_0} \nu$  with respect to  $m_0$  is uniformly bounded in  $i$ . We write first

$$\frac{dF_*^k \nu}{dm}(x) = \frac{dF_*^k \nu}{dF_*^k m_0} \frac{dF_*^k m_0}{dm}(x) = \sum_{x_0 \in F^{-k}(x)} \frac{\frac{d\nu}{dm_0}(x_0)}{JF^n(x_0)} \leq \left\| \frac{d\nu}{dm_0} \right\|_\infty \left\| \frac{dF_*^k m_0}{dm} \right\|_\infty.$$

The second factor is bounded by  $M_0$  from Lemma A.3. Let us estimate the first one. Note that,  $\frac{d\nu}{dm_0}(x) = \frac{1}{JF^{\tau_i-1} x_0}$ , where  $x_0 = (F^{\tau_i-1} | \Omega)^{-1}(x)$ . Corollary A.5 implies that  $\|d\nu/dm\|_\infty \leq (1+D)/m_0(\Delta_0)$  and so we get

$$\bar{m}_0\{\tau_{i+1} - \tau_i > n_0 + n | \Gamma\} = \bar{m}_0\{\tau_{i+1} - \tau_i - n_0 > n | \Gamma\} \leq M_0 \frac{1+D}{m_0(\Delta_0)} m_0\{\hat{R} > n\}.$$

For  $n > n_0$ , and using the definition of  $\hat{R}$ , we can write this is

$$\bar{m}_0\{\tau_{i+1} - \tau_i > n | \Gamma\} \leq M_0 \frac{1+D}{m_0(\Delta_0)} m_0\{\hat{R} > n - n_0\} = M_0 \frac{1+D}{m_0(\Delta_0)} \sum_{i \geq n-n_0} m_0\{R > i\}.$$

For  $i$  odd the calculation is exactly the same and we get

$$\bar{m}_0\{\tau_{i+1} - \tau_i > n | \Gamma\} \leq M'_0 \frac{1+D'}{m'_0(\Delta'_0)} m'_0 \sum_{i \geq n-n_0} m'_0\{R' > i\}.$$

Letting  $K_0 = \max \left\{ \frac{M_0(D+1)}{m_0(\Delta_0)}, \frac{M'_0(D'+1)}{m'_0(\Delta'_0)} \right\}$  and using (1.2) we get the assertion.  $\square$

## 5. RATES OF DECAY

We now fix some arbitrary  $n \geq 1$  and estimate  $\bar{m}_0\{T > n\}$ . Letting  $\tau_0 = 0$  we write

$$(5.1) \quad \bar{m}_0\{T > n\} = \sum_{i \geq 1} \bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\}.$$

We will estimate the right hand side of (5.1) using different arguments depending on whether the decay of  $\mathcal{M}_n$  is exponential, stretched exponential or polynomial.

**5.1. Polynomial case.** We suppose that  $\mathcal{M}_n = \mathcal{O}(n^{-\alpha})$  for some  $\alpha > 2$  and prove that

$$(5.2) \quad \bar{m}_0\{T > n\} = \mathcal{O}(n^{1-\alpha})$$

Let  $K = 2 \max\{\bar{m}_0(\bar{\Delta}_0), K_0, m_0(\Delta_0), m'(\Delta'_0)\}$  and  $\varepsilon_0 > 0$  given by Proposition 4.1. We start with the following somewhat unwieldy estimate.

**Proposition 5.1.** *For any  $n \in \mathbb{N}$*

$$(5.3) \quad \bar{m}_0\{T > n\} \leq K \sum_{i \leq \frac{1}{2} \left\lfloor \frac{n}{n_0} \right\rfloor} i(1 - \varepsilon_0)^{i-3} \bar{\mathcal{M}}_{\left\lfloor \frac{n}{i} \right\rfloor - n_0} + \bar{m}_0(\bar{\Delta}_0)(1 - \varepsilon_0)^{\frac{1}{2} \left\lfloor \frac{n}{n_0} \right\rfloor - 1}.$$

Before proving Proposition 5.1 we show how it implies (5.2).

*Proof of Theorem in the polynomial case assuming Proposition 5.1.* By the definition of  $\bar{\mathcal{M}}_{\left\lfloor \frac{n}{i} \right\rfloor - n_0}$  in (4.1) and the assumption on the asymptotics of  $\mathcal{M}_n$ , there exists a constant

$C > 0$  such that for every  $i \leq \frac{1}{2} \left\lfloor \frac{n}{n_0} \right\rfloor$  we have

$$\begin{aligned} \bar{\mathcal{M}}_{\left\lfloor \frac{n}{i} \right\rfloor - n_0} &= \sum_{j \geq \left\lfloor \frac{n}{i} \right\rfloor - n_0} \mathcal{M}_j \leq C \sum_{j \geq \left\lfloor \frac{n}{i} \right\rfloor - n_0} j^\alpha \leq C \int_{\left\lfloor \frac{n}{i} \right\rfloor - n_0}^{\infty} x^{-\alpha} dx \\ &\leq C \frac{i^{\alpha-1}}{n^{\alpha-1}} \left( \frac{n}{n - i(n_0 + 1)} \right)^{\alpha-1} \leq C \frac{i^{\alpha-1}}{n^{\alpha-1}}. \end{aligned}$$

Substituting this into the statement of Proposition 5.1 we get

$$(5.4) \quad \bar{m}_0\{T > n\} \leq \frac{KC}{n^{\alpha-1}} \sum_{i \leq \frac{1}{2} \left\lfloor \frac{n}{n_0} \right\rfloor} i^\alpha (1 - \varepsilon_0)^{i-3} + \bar{m}_0(\bar{\Delta}_0)(1 - \varepsilon_0)^{\frac{1}{2} \left\lfloor \frac{n}{n_0} \right\rfloor - 1}.$$

Since the series  $\sum_{i=1}^{\infty} (1 - \varepsilon_0)^{i-3} i^\alpha$  is convergent and the second term in (5.4) is exponentially small in  $n$ , we get (5.2) and thus the statement of the Theorem in the polynomial case.  $\square$

The proof of Proposition 5.1 will be broken into several lemmas. Note that

$$\sum_{i > \frac{1}{2} \left\lfloor \frac{n}{n_0} \right\rfloor} \bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} \leq \bar{m}_0\{T > n; \tau_{\frac{1}{2} \left\lfloor \frac{n}{n_0} \right\rfloor} \leq n\},$$

which together with (5.1) implies

$$(5.5) \quad \bar{m}_0\{T > n\} \leq \sum_{i \leq \frac{1}{2} \lfloor \frac{n}{n_0} \rfloor} \bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} + \bar{m}_0\{T > n; \tau_{\frac{1}{2} \lfloor \frac{n}{n_0} \rfloor} \leq n\}.$$

First we estimate the second summand of (5.5).

**Lemma 5.2.** For every  $n > 2n_0$  and for  $\varepsilon_0 > 0$  as in Proposition 4.1

$$\bar{m}_0\{T > n; \tau_{\frac{1}{2} \lfloor \frac{n}{n_0} \rfloor} \leq n\} \leq \bar{m}_0(\bar{\Delta}_0)(1 - \varepsilon_0)^{\frac{1}{2} \lfloor \frac{n}{n_0} \rfloor - 1}.$$

*Proof.* Since  $T > n > \tau_{i-1}$  we have

$$\begin{aligned} & \bar{m}_0\{T > n; \tau_{\frac{1}{2} \lfloor \frac{n}{n_0} \rfloor} \leq n\} \leq \bar{m}_0\{T > \tau_{\frac{1}{2} \lfloor \frac{n}{n_0} \rfloor}\} = \\ & \bar{m}_0\{T > \tau_1\} \bar{m}_0\{T > \tau_2 \mid T > \tau_1\} \dots \bar{m}_0\{T > \tau_{\frac{1}{2} \lfloor \frac{n}{n_0} \rfloor} \mid T > \tau_{\frac{1}{2} \lfloor \frac{n}{n_0} \rfloor - 1}\}. \end{aligned}$$

Notice that  $\bar{m}_0\{T > \tau_1\} \leq \bar{m}_0(\bar{\Delta}_0)$ . The first item of Proposition 4.1 implies that each of the other terms is less than  $1 - \varepsilon_0$ . Substituting these into the above equation finishes the proof.  $\square$

Now, we begin estimating the first summand of (5.5). Start with the cases  $i = 1, 2$ .

**Lemma 5.3.** For  $i = 1, 2$  and every  $n > n_0$  we have

$$\bar{m}_0\{T > \tau_{i-1}; \tau_{i-1} \leq n < \tau_i\} \leq K \bar{\mathcal{M}}_{\lfloor \frac{n}{2} \rfloor - n_0}.$$

*Proof.* For  $i = 1$  we take advantage of the fact that  $\tau_1$  depends only on the first coordinate. Then we have

$$\begin{aligned} \bar{m}_0\{T > 0; \tau_1 > n\} &= m_0\{\tau_1 > n\} m'_0(\Delta'_0) \leq m'_0(\Delta'_0) m\{\hat{R} > n - n_0\} \\ &\leq m'_0(\Delta'_0) \sum_{j \geq n - n_0} m_0\{R > j\} \leq m'_0(\Delta'_0) \bar{\mathcal{M}}_{n - n_0}. \end{aligned}$$

which proves the statement in this case by the definition of  $K$  and using the fact that  $\bar{\mathcal{M}}_k$  is monotone decreasing in  $k$ . For  $i = 2$  we have

$$\bar{m}_0\{T > \tau_1; \tau_1 \leq n < \tau_2\} \leq \bar{m}_0\{\tau_2 > n\} \leq \bar{m}_0\{\tau_2 - \tau_1 \geq \frac{n}{2}\} + \bar{m}_0\{\tau_1 \geq \frac{n}{2}\}.$$

From the second item of Proposition 4.1 we have

$$\bar{m}_0\{\tau_2 - \tau_1 \geq \frac{n}{2}\} \leq K_0 \bar{\mathcal{M}}_{\lfloor \frac{n}{2} \rfloor - n_0}.$$

The second item is estimated as in the case  $i = 1$  and so we get

$$\bar{m}_0\{T > \tau_1; \tau_1 \leq n < \tau_2\} \leq K_0 \bar{\mathcal{M}}_{\lfloor \frac{n}{2} \rfloor - n_0} + m'_0(\Delta'_0) \bar{\mathcal{M}}_{\lfloor \frac{n}{2} \rfloor - n_0} \leq K \bar{\mathcal{M}}_{\lfloor \frac{n}{2} \rfloor - n_0}.$$

which completes the proof in this case also.  $\square$

We now consider the general case.

**Lemma 5.4.** For each  $i \geq 3$

$$\bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} \leq \sum_{j=1}^i K(1 - \varepsilon_0)^{i-3} \bar{\mathcal{M}}_{\lfloor \frac{n}{i} \rfloor - n_0}.$$

*Proof.* Since  $T > n \geq \tau_{i-1}$  we have

$$\bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} \leq \bar{m}_0\{T > \tau_{i-1}; n < \tau_i\}.$$

Moreover, from  $\tau_i = \tau_i - \tau_{i-1} + \tau_{i-1} - \tau_{i-2} + \dots + \tau_1 - \tau_0 > n$  we obtain that there is at least one  $j \in [1, i]$  such that  $\tau_j - \tau_{j-1} > \frac{n}{i}$  and therefore

$$\bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} \leq \sum_{j=1}^i \bar{m}_0\{T > \tau_{i-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\}.$$

For each  $i, j \geq 3$  we write

$$(5.6) \quad \bar{m}_0\{T > \tau_{i-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\} = Y_1 \cdot Y_2 \cdot Y_3$$

where

$$Y_1 := \bar{m}_0\{T > \tau_{j-1}; \tau_j - \tau_{j-1} > \frac{n}{i} | T > \tau_{j-2}\},$$

$$Y_2 := \bar{m}_0\{T > \tau_1\} \bar{m}_0\{T > \tau_2 | T > \tau_1\} \dots \bar{m}_0\{T > \tau_{j-2} | T > \tau_{j-3}\},$$

$$Y_3 := \bar{m}_0\{T > \tau_j | T > \tau_{j-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\} \dots \bar{m}_0\{T > \tau_{i-1} | T > \tau_{i-2}; \tau_j - \tau_{j-1} > \frac{n}{i}\}.$$

By the second item of Proposition 4.1 we have

$$(5.7) \quad Y_2 \leq \bar{m}_0(\bar{\Delta}_0)(1 - \varepsilon_0)^{j-3}.$$

For the first term, note that

$$Y_1 := \bar{m}_0\{T > \tau_{j-1}; \tau_j - \tau_{j-1} > \frac{n}{i} | T > \tau_{j-2}\} \leq \bar{m}_0\{\tau_j - \tau_{j-1} > \frac{n}{i} | T > \tau_{j-2}\}.$$

By construction  $\{T > \tau_{j-2}\}$  can be written as a union of elements of  $\xi_{j-1}$  and so, by the second item of Proposition 4.1,

$$(5.8) \quad Y_1 \leq K_0 \bar{\mathcal{M}}_{\lfloor \frac{n}{i} \rfloor - n_0}.$$

For the third term, since  $\tau_j$  and  $\tau_{j-1}$  are constant on the elements of  $\xi_j$ , if  $\tau_j(\bar{x}) - \tau_{j-1}(\bar{x}) > \frac{n}{i}$  for some point  $\bar{x}$ , then it holds on  $\xi_j(\bar{x})$ . By construction, for  $k \geq j$  the partition  $\xi_k$  is finer than  $\xi_j$  and  $\{T > \tau_{k-1}\}$  can be written as a union of elements of  $\xi_k$ . Hence  $\{T > \tau_{k-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\}$  can be covered with elements of  $\xi_k$ . Using the first item of Proposition 4.1 in each partition element gives

$$\bar{m}_0\{T > \tau_k | T > \tau_{k-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\} \leq 1 - \varepsilon_0.$$

This immediately gives

$$(5.9) \quad Y_3 \leq (1 - \varepsilon_0)^{i-j}.$$



Substituting (5.7), (5.8), (5.9) into (5.6) we get the assertion of Lemma 5.5. For the case  $i \geq 3$  and  $j < 3$  proof will be the same but only without  $Y_2$ .  $\square$

Notice that substituting the estimates in the statements of Lemmas 5.3 and 5.4 into (5.5) gives the statement in Proposition 5.1.

**5.2. Super polynomial cases.** In this subsection we give the proof of the tail estimates for exponential and stretched exponential cases. Let

$$A(i) = \{\mathbf{k} = (k_1, \dots, k_{i-1}) \in \mathbb{N}^{i-1} : \sum_j k_j \leq n, k_j \geq n_0, j = 1, \dots, i-1\},$$

$$(5.10) \quad \bar{\mathcal{M}}(i, n) = \max_{\mathbf{k} \in A(i)} \bar{\mathcal{M}}_{n - \sum_j k_j - n_0} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j - n_0}.$$

*Remark 3.* It is known fact (see for example [45]) that the cardinality  $\text{card} A(i)$  of  $A(i)$  is bounded above by  $\binom{n+i-n_0}{i-1}$ .

Let  $\delta$  be a sufficiently small number, which will be specified later. We first prove the following technical statement from which both exponential and stretched exponential cases will follow.

**Proposition 5.5.** *For sufficiently large  $n$  and any  $\theta' \in (0, 1]$  we have*

$$\bar{m}_0\{T > n\} \leq \sum_{i \leq [\delta n^{\theta'}]} \binom{n+i-n_0}{i-1} K_0^i \bar{\mathcal{M}}(i, n) + \bar{m}(\bar{\Delta}_0)(1 - \varepsilon_0)^{[\delta n^{\theta'}]-1}.$$

To prove Proposition 5.5 we first write, as in (5.5),

$$(5.11) \quad \bar{m}_0\{T > n\} \leq \sum_{i \leq [\delta n^{\theta'}]} \bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} + \bar{m}_0\{T > n; \tau_{[\delta n^{\theta'}]} \leq n\}.$$

As in polynomial case, we start by estimating the second summand of (5.11).

**Lemma 5.6.** For sufficiently large  $n$  we have

$$\bar{m}_0\{T > n; \tau_{[\delta n^{\theta'}]} \leq n\} \leq \bar{m}(\bar{\Delta}_0)(1 - \varepsilon_0)^{[\delta n^{\theta'}]-1}.$$

*Proof.* The proof is identical to the proof of Lemma 5.2.  $\square$

To estimate the second summand of (5.11), first we fix  $i$  and prove the following

**Lemma 5.7.** For sufficiently large  $n$ , for every  $i \leq [\delta n^{\theta'}]$  we have

$$\bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} \leq \bar{m}(\bar{\Delta}_0) K_0^i \text{card} A(i) \bar{\mathcal{M}}(i, n).$$

*Proof.* For any  $\bar{x} \in \bar{\Delta}_0$  with  $\tau_{i-1}(\bar{x}) \leq n < \tau_i(\bar{x})$  there is  $(k_1, \dots, k_{i-1}) \in A(i)$  such that  $k_j = \tau_j(\bar{x}) - \tau_{j-1}(\bar{x})$  for  $j = 1, \dots, i-1$  and  $\tau_i(\bar{x}) - \tau_{i-1}(\bar{x}) > n - \sum_j k_j$ . Now, for every  $\mathbf{k} = (k_1, \dots, k_{i-1}) \in A(i)$  let

$$(5.12) \quad P(\mathbf{k}, i) = \bigcap_{j=1}^{i-1} \{\bar{x} \in \bar{\Delta} : \tau_j(\bar{x}) - \tau_{j-1}(\bar{x}) = k_j\}$$

and

$$(5.13) \quad Q(\mathbf{k}, i) = P(\mathbf{k}, i) \cap \{\bar{x} \in \bar{\Delta} : \tau_i(\bar{x}) - \tau_{i-1}(\bar{x}) > n - \sum_j k_j\}.$$

Using the above observation and notations we can write

$$(5.14) \quad \bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} \leq \bar{m}_0\{\tau_{i-1} \leq n < \tau_i\} = \sum_{\mathbf{k} \in A(i)} \bar{m}_0\{Q(\mathbf{k}, i)\}.$$

Notice that for each  $i$  and each  $\mathbf{k} \in A(i)$  we have

$$(5.15) \quad \begin{aligned} \bar{m}_0\{Q(\mathbf{k}, i)\} &= \bar{m}_0(\bar{\Delta}_0) \bar{m}_0\{\tau_1 = k_1 | \bar{\Delta}_0\} \bar{m}_0\{\tau_2 - \tau_1 = k_2 | P(\mathbf{k}, 2)\} \dots \\ &\dots \bar{m}_0\{\tau_i - \tau_{i-1} > n - \sum k_j | P(\mathbf{k}, i)\}. \end{aligned}$$

Since for any  $j$ , the set  $\{P(\mathbf{k}, j)\}$  is a union of elements of  $\xi_j$ , from the second item of Proposition 4.1 we get

$$\bar{m}_0\{\tau_j - \tau_{j-1} = k_j | P(\mathbf{k}, j)\} \leq K_0 \bar{\mathcal{M}}_{k_j - n_0}$$

for each  $1 \leq j \leq i-1$ , and

$$\bar{m}_0\{\tau_i - \tau_{i-1} > n - n_0 - \sum k_j | P(\mathbf{k}, i)\} \leq K_0 \bar{\mathcal{M}}_{n - n_0 - \sum k_j}.$$

Substituting this into (5.15) and using the definition of  $\bar{\mathcal{M}}(i, n)$  in (5.10), we obtain

$$\bar{m}_0\{Q(\mathbf{k}, i)\} \leq \bar{m}_0(\bar{\Delta}_0) K_0^i \bar{\mathcal{M}}_{n - n_0 - \sum k_j} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j - n_0} \leq \bar{m}_0(\bar{\Delta}_0) K_0^i \bar{\mathcal{M}}(i, n).$$

Substituting this into (5.14) completes the proof of Lemma 5.7.  $\square$

*Proof of Proposition 5.5.* To prove the Proposition we just substitute the statements in the two Lemmas above into (5.11) and use the upper bound for the cardinality of  $A(i)$ , see Remark (3).  $\square$

*Proof of the Theorem in the stretched exponential case.* Here we consider the case  $\mathcal{M}_n = \mathcal{O}(e^{-\tau n^\theta})$  for some  $\tau, \theta > 0$  and we show that

$$(5.16) \quad \bar{m}_0\{T > n\} \leq C e^{-\tau' n^{\theta'}}$$

for some  $C_1, \tau' > 0$  and  $\theta > \theta' > 0$ . We start with the following basic estimate

**Lemma 5.8.** For sufficiently large  $n$ , every  $i \leq [\delta n^{\theta'}]$  and every  $\mathbf{k} \in A(i)$  we have

$$\bar{\mathcal{M}}_{n-\sum_j k_j-n_0} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j-n_0} \leq C_1^i e^{-\tau(n-in_0)^\theta} n^{i(1-\theta)}.$$

The statement and proof of Lemma 5.8 depend only on some relatively standard but non-trivial estimates concerning the tails of sequences which decay at a stretched exponential rate. In order to simplify the exposition we give the proof in the appendix. From Lemma 5.8 and the definition of  $\bar{\mathcal{M}}(i, n)$  we obtain

$$(5.17) \quad \bar{\mathcal{M}}(i, n) \leq C_1^i e^{-\tau(n-in_0)^\theta} n^{i(1-\theta)}.$$

On the other hand combining Pascal's rule with Stirling's formula we get

$$(5.18) \quad \binom{n+i-n_0}{i-1} \leq \binom{n}{[\delta n^{\theta'}]} \leq C_2 e^{\varepsilon n^{\theta'} \log n} < C_2 e^{\varepsilon n^\theta}$$

for  $\varepsilon > 0$  such that  $\varepsilon \rightarrow 0$  as  $\delta \rightarrow 0$ . Substituting inequalities (5.17) and (5.18) into Proposition 5.5 we get

$$(5.19) \quad \bar{m}\{T > n\} \leq \sum_{i \leq [\delta n^{\theta'}]} C_2 e^{\varepsilon n^\theta} K_0^i C_1^i e^{-\tau(n-in_0)^\theta} n^{i(1-\theta)} + \bar{m}(\bar{\Delta}_0)(1-\varepsilon_0)^{[\delta n^{\theta'}]-1}.$$

The second term of (5.19) is of order  $e^{n^{\theta'} \delta \log(1-\varepsilon_0)}$ . Hence, it remains to prove similar asymptotics for the first summand. Notice first of all that the terms in the sum are monotone increasing in  $i$ . Therefore denoting these terms by  $a_i$  we have that

$$\sum_{i \leq [\delta n^{\theta'}]} a_i \leq a_{[\delta n^{\theta'}]} [\delta n^{\theta'}] = e^{\varepsilon n^\theta} (K_0 C_1)^{[\delta n^{\theta'}]} e^{-\tau(n-[\delta n^{\theta'}]n_0)^\theta} n^{[\delta n^{\theta'}](1-\theta)} [\delta n^{\theta'}].$$

Writing the right hand side in exponential form, we have

$$a_{[\delta n^{\theta'}]} [\delta n^{\theta'}] = \exp(\varepsilon n^\theta + \delta n^{\theta'} \log(K_0 C_1) - \tau(n - n_0 [\delta n^{\theta'}])^\theta + (1-\theta)[\delta n^{\theta'}] \log n + \log(\delta n^{\theta'})).$$

Factoring out  $n^{\theta'}$  and using the general inequality  $a^\theta - b^\theta \leq (a-b)^\theta$  the exponent is  $\leq$

$$n^{\theta'} \left\{ (\varepsilon - \tau) n^{\theta-\theta'} + \delta \log(K_0 C_1) + \tau(\delta n_0)^\theta n^{(\theta-1)\theta'} + (1-\theta)\delta \log n + n^{-\theta'} \log(\delta n^{\theta'}) \right\}.$$

Since this last expression is clearly decreasing in  $\delta$  and negative for  $\delta > 0$  sufficiently small we get the stretched exponential bound as required.  $\square$

*Proof of the Theorem in the exponential case.* Here we consider the case  $\mathcal{M}_n = \mathcal{O}(e^{-\tau n})$  for some  $\tau > 0$  and we show that

$$(5.20) \quad \bar{m}_0\{T > n\} \leq C e^{-\tau' n}$$

for some  $\tau' > 0$ .

**Lemma 5.9.** For all  $i \leq [\delta n]$  and every  $\mathbf{k} \in A(i)$  we have

$$\bar{\mathcal{M}}_{n-\sum_j k_j-n_0} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j-n_0} \leq C^i e^{-\tau(n-in_0)}.$$

*Proof.* Since  $\mathcal{M}_n$  is decaying exponentially fast, there exists  $C' > 0$  such that for any  $j = 1, \dots, i-1$  we have

$$\bar{\mathcal{M}}_{k_j-n_0} = \sum_{\kappa \geq k_j-n_0} \mathcal{M}_\kappa \leq C' \sum_{\kappa \geq k_j-n_0} e^{-\tau\kappa} = \frac{C'}{1-e^{-\tau}} e^{-\tau(k_j-n_0)}$$

and similarly

$$\bar{\mathcal{M}}_{n-n_0-\sum k_j} \leq \frac{C'}{1-e^{-\tau}} e^{-\tau(n-n_0-\sum k_j)}.$$

Letting  $C := C'/(1-e^{-\tau})$  and substituting these estimates into the expression in the lemma finishes the proof.  $\square$

By Lemma 5.9 and the definition of  $\bar{\mathcal{M}}(i, n)$  we obtain

$$(5.21) \quad \bar{\mathcal{M}}(i, n) \leq C^i e^{-\tau(n-in_0)}.$$

On the other hand, since  $i \leq [\delta n]$ , there exists a uniform constant  $C_1$  such that

$$(5.22) \quad \binom{n+i-n_0}{i-1} \leq \binom{n}{[\delta n]} \leq C_1 e^{\varepsilon n}$$

for  $\varepsilon > 0$  such that  $\varepsilon \rightarrow 0$  as  $\delta \rightarrow 0$ . Choosing  $\theta' = 1$  in Proposition 5.5 and substituting inequalities (5.21) and (5.22) we obtain

$$(5.23) \quad \bar{m}_0\{T > n\} \geq \sum_{i \leq [\delta n]} C_1 e^{\varepsilon n} (K_0 C)^i e^{-\tau(n-in_0)} + \bar{m}(\Delta_0)(1-\varepsilon_0)^{[\delta n]-1}.$$

Choose  $\delta$  small enough so that  $\tau'(\delta) = \tau(1-\delta n_0) - \delta \log(K_0 C) - \varepsilon > 0$ . This is possible, since as  $\delta \rightarrow 0$  we have  $\tau'(\delta) \rightarrow \tau$ . Notice, that each term of the first summand in (5.23) can be written in an exponential form with exponent

$$\varepsilon n + i \log(K_0 C) - \tau(n - in_0) = -\tau'(\delta)n + \tau n_0(i - \delta n) + (i - \delta n) \log(K_0 C).$$

Hence, we can write the first summand of (5.23) as

$$(5.24) \quad C_1 e^{-\tau'(\delta)n} \sum_{i \leq [\delta n]} e^{(i-\delta n)(\log(K_0 C) + \tau n_0)} \leq C_1 e^{-\tau'(\delta)n} \sum_{j \geq 1} e^{-j(\log(K_0 C) + \tau n_0)}.$$

Notice that the series in the right hand side of (5.24) is convergent. Therefore, by choosing  $\tau' = \min\{\tau'(\delta), \delta \log(1-\varepsilon_0)\}$  we get (5.20).  $\square$

## 6. PROOF IN THE NON-INVERTIBLE SETTING

In this section we give the proof of the Theorem in the non-invertible setting. We will prove each of the required properties (G1)-(G5) in separate subsections, thus completing the proof of the existence of a tower for the product system  $f : M \rightarrow M$ . The tail estimates obtained above in the exponential, stretched exponential and polynomial case, then complete the proof.

**(G1) Markov property.** In Section 3 we carried out a the construction of the collection  $\bar{\eta}$  of subsets of  $\bar{\Delta}_0$  which, by the first item of Proposition 4.1 forms a partition of  $\bar{\Delta}_0 \bmod 0$ . In particular we have an induced map  $\bar{F}^T : \bar{\Delta}_0 \rightarrow \bar{\Delta}_0$  which, using the canonical identification between  $\bar{\Delta}_0$  as the base of the tower and as a subset of the ambient product manifold  $M$ , corresponds to an induced map  $\bar{f}^T : \bar{\Delta}_0 \rightarrow \bar{\Delta}_0$  where  $\bar{f} : M \rightarrow M$  is the product map on  $M$ . The construction of  $\bar{\eta}$  and the induced map implies that  $\bar{f}^T$  satisfies the Markov property (G1).

**(G2) Uniform Expansion.** The uniform expansivity condition follows immediately from the fact that it holds by assumption for the individual components and that the return time  $T$  is a sum of return times for each of the individual components. Then

$$\|(D\bar{f}^T)^{-1}(x)\| \leq \max\{\|(Df^T)^{-1}(x)\|, \|(Df'^T)^{-1}(x)\|\} \leq \beta.$$

**(G3) Bounded distortion.** Recall first of all from the definition of Gibbs-Markov-Young tower the notion of *separation time* and let  $s, s'$  denote the separation time of points in  $\Delta_0, \Delta'_0$  with respect to the partitions  $\eta, \eta'$  respectively. We let  $D, \beta, D', \beta'$  be the distortion constants as in (2.1) for  $f$  and  $f'$  respectively and let

$$\bar{D} = \max\left\{\frac{D\beta}{1-\beta}, \frac{D'\beta'}{1-\beta'}\right\}, \quad \bar{\beta} = \max\{\beta, \beta'\}.$$

Now let  $\bar{s}$  denote the separation time of points in  $\bar{\Delta}_0$  with respect to the product map  $\bar{f}$  and the partition  $\bar{\eta}$  constructed above. Then, using the usual identification of  $\bar{\Delta}_0$  with the subset of manifold, to prove the bounded distortion condition (G3) it is sufficient to show that for any  $\Gamma \in \bar{\eta}$  and  $\bar{x}, \bar{y} \in \Gamma$  such that  $\bar{s}(\bar{x}, \bar{y}) < \infty$  we have

$$\left| \log \frac{\det D\bar{f}^T(\bar{x})}{\det D\bar{f}^T(\bar{y})} \right| \leq \bar{D}\bar{\beta}^{\bar{s}(\bar{f}^T(\bar{x}), \bar{f}^T(\bar{y}))}.$$

To prove this, note first that, by the property of Jacobian and absolute value we have

$$(6.1) \quad \left| \log \frac{\det D\bar{f}^T(\bar{x})}{\det D\bar{f}^T(\bar{y})} \right| \leq \left| \log \frac{\det f^T(x)}{\det Df^T(y)} \right| + \left| \log \frac{\det Df'^T(x')}{\det Df'^T(y')} \right|$$

for all  $\bar{x} = (x, x'), \bar{y} = (y, y') \in \Gamma$ . Moreover, notice that since the determinants are all calculated at return times  $T$ , we can use the identification  $\bar{f}^T = \bar{F}^T$  and reformulate the above expressions in terms of the Jacobians  $JF^T$ .

For the first term, notice that the simultaneous return time  $T$  can be written as a sum of return times to  $\Delta_0$ . Without loss of generality, assume  $T = R_1 + \dots + R_k$ , for some  $k$ . For any  $x \in \Delta_0$ , let  $x_0 = x$  and  $x_j = F^{R_1+\dots+R_j}(x)$  where  $j = 1, \dots, k-1$ . Then

$$JF^T(x) = JF^{R_1}(x)JF^{R_2}(x_1)JF^{R_3}(x_2)\dots JF^{R_k}(x_{k-1}).$$

Since  $F^T(\pi\Gamma) = \Delta_0$  we have  $F^{R_1}(\pi\Gamma) \subset \eta(x_1)$ , ...,  $F^{R_1+\dots+R_{k-1}}(\pi\Gamma) \subset \eta(x_{k-1})$ . Hence, for the points  $x, y \in \pi\Gamma$  the sequences  $x_j$  and  $y_j$  belong to the same element of  $\eta$  for all  $j = 0, \dots, k-1$ , which implies

$$s(F^{R_{j+1}}(x_j), F^{R_{j+1}}(y_j)) = s(F^T(x), F^T(y)) + R_{j+2} + \dots + R_k.$$

Hence,

$$\begin{aligned} \left| \log \frac{JF^T(x)}{JF^T(y)} \right| &\leq \sum_{j=1}^k \left| \log \frac{JF^{R_j}(x)}{JF^{R_j}(y)} \right| \leq \sum_{j=1}^k \left| \frac{JF^{R_j}(x)}{JF^{R_j}(y)} - 1 \right| \\ &\leq \sum_{j=0}^k D\beta^{s(F^{R_{j+1}}(x_j), F^{R_{j+1}}(y_j))} \leq D\beta^{s(F^T(x), F^T(y))} \sum_{j=0}^k \beta^{R_{j+2}+\dots+R_k} \\ &\leq D\beta^{s(F^T(x), F^T(y))} \sum_{j=0}^{\infty} \beta_F^j \leq D \frac{\beta}{1-\beta} \beta^{s(F^T(x), F^T(y))}. \end{aligned}$$

The second summand is estimated similarly and we get

$$\left| \log \frac{JF'^T(x')}{JF'^T(y')} \right| \leq D' \frac{\beta'}{1-\beta'} \beta'^{s'(F'^T(x'), F'^T(y'))}.$$

Using the fact that for any  $w, z \in \bar{\Delta}_0$  we have  $\bar{s}(w, z) \leq \min\{s(\pi w, \pi z), s'(\pi' w, \pi' z)\}$  we obtain the required bound.

**(G4) Integrability.** Follows immediately from the tail estimates obtained above.

**6.1. Aperiodicity.** As mentioned above, conditions (G1)-(G4) imply the existence of an ergodic  $\bar{f}$ -invariant probability measure  $\bar{\mu}$ . Moreover it is known by standard results that this measure is mixing if and only if the aperiodicity conditions is satisfied. Thus it is sufficient to show that  $\bar{\mu}$  is mixing. To see this, let  $\mu, \mu'$  be the invariant, mixing, probability measures associated to the maps  $f, f'$  as introduced in Section 3.2. Then the measure  $\mu \times \mu'$  is invariant and mixing for the product map  $\bar{f}$  and thus it is sufficient to show that  $\bar{\mu} = \mu \times \mu'$  to imply that  $\bar{\mu}$  is mixing. This follows again by standard uniqueness arguments. Indeed, both  $\bar{\mu}$  and  $\mu \times \mu'$  are ergodic and equivalent to the reference measure, at least on the set  $\bar{\Delta}_0$ . Thus, by Birkhoff's ergodic Theorem, for any integrable function, their time averages converge to the same limit and so  $\int \varphi d\bar{\mu} = \int \varphi d\mu \times \mu'$  implying that  $\bar{\mu} = \mu'$ .

## 7. PROOF IN THE INVERTIBLE SETTING

Here we give proof of the Theorem in the invertible setting. We start with the definition of Young towers in this setting, following [5] which generalizes the definition of [49] (see Remarks 2.3-2.5 from [5]). We change the notation slightly to distinguish this case from non-invertible case. Let  $M$  be a Riemannian manifold. If  $\gamma \subset M$  is a submanifold, then  $m_\gamma$  denotes the restriction of the Riemannian volume to  $\gamma$ . Consider  $f : M \rightarrow M$  such that  $f : M \setminus \mathcal{C} \rightarrow M \setminus f(\mathcal{C})$  is a  $C^{1+\varepsilon}$  diffeomorphism on each connected component of  $M \setminus \mathcal{C}$  for some  $\mathcal{C} \subset M$  with the following properties.

- (A1) There exists  $\Lambda \subset M$  with hyperbolic product structure, i.e. there are families of stable and unstable manifolds  $\Gamma^s = \{\gamma^s\}$  and  $\Gamma^u = \{\gamma^u\}$  such that  $\Lambda = (\cup \gamma^s) \cap (\cup \gamma^u)$ ;  $\dim \gamma^s + \dim \gamma^u = \dim M$ ; each  $\gamma^s$  meets each  $\gamma^u$  at a unique point; stable and unstable manifolds are transversal with angles bounded away from 0;  $m_{\gamma^u}(\gamma^u \cap \Lambda) > 0$  for any  $\gamma^u$ .

Let  $\Gamma^s$  and  $\Gamma^u$  be the defining families of  $\Lambda$ . A subset  $\Lambda_0 \subset \Lambda$  is called  $s$ -subset if  $\Lambda_0$  also has a hyperbolic structure and its defining families can be chosen as  $\Gamma^u$  and  $\Gamma_0^s \subset \Gamma^s$ . Similarly, we define  $u$ -subsets. For  $x \in \Lambda$  let  $\gamma^\theta(x)$  denote the element of  $\Gamma^\theta$  containing  $x$ , where  $\theta = u, s$ .

- (A2) There are pairwise disjoint  $s$ -subsets  $\Lambda_1, \Lambda_2, \dots \subset \Lambda$  such that  $m_{\gamma^u}((\Lambda \setminus \cup \Lambda_i) \cap \gamma^u) = 0$  on each  $\gamma^u$  and for each  $\Lambda_i$ ,  $i \in \mathbb{N}$  there is  $R_i$  such that  $f^{R_i}(\Lambda_i)$  is  $u$ -subset;  $f^{R_i}(\gamma^s(x)) \subset \gamma^s(f^{R_i}(x))$  and  $f^{R_i}(\gamma^u(x)) \supset \gamma^u(f^{R_i}(x))$  for any  $x \in \Lambda_i$ .
- (A3) There exist constants  $C \geq 1$  and  $\beta \in (0, 1)$  such that  $\text{dist}(f^n(x), f^n(y)) \leq C\beta^n$ , for all  $y \in \gamma^s(x)$  and  $n \geq 0$ ;
- (A4) Regularity of the stable foliation: given  $\gamma, \gamma' \in \Gamma^u$  define  $\Theta : \gamma' \cap \Lambda \rightarrow \gamma \cap \Lambda$  by  $\Theta(x) = \gamma^s(x) \cap \gamma$ . Then
- (a)  $\Theta$  is absolutely continuous and

$$u(x) := \frac{d(\Theta_* m_{\gamma'})}{dm_\gamma}(x) = \prod_{i=0}^{\infty} \frac{\det Df^u(f^i(x))}{\det Df^u(f^i(\Theta^{-1}(x)))};$$

- (b) There exists  $C > 0$  and  $\beta < 1$  such that, letting the separation time  $s(x, y)$  be the smallest  $k$  where  $(f^R)^k(x)$  and  $(f^R)^k(y)$  lie in different partition elements, we have

$$\log \frac{u(x)}{u(y)} \leq C\beta^{s(x,y)} \quad \text{for } x, y \in \gamma' \cap \Lambda.$$

- (A5) Bounded distortion: for  $\gamma \in \Gamma^u$  and  $x, y \in \Lambda \cap \gamma$

$$\log \frac{\det D(f^R)^u(x)}{\det D(f^R)^u(y)} \leq C\beta^{s(f^R(x), f^R(y))}.$$

- (A6)  $\int R dm_0 < \infty$ , where  $m_0$  is the restriction of Lebesgue measure to  $\Lambda$ .

- (A7)  $\gcd\{R_i\} = 1$ .

Given such a structure we can define Young-tower as we define  $\Delta$  in the non-invertible case. This time we denote the tower by  $\mathcal{T}$ . Let  $F_i : \mathcal{T}_i \rightarrow \mathcal{T}_i$  be the two towers corresponding to maps  $f_i : M_i \rightarrow M_i$ ,  $i = 1, 2$  as in the statement of the Theorem. Then, from conditions (A1)-(A7) we can obtain GMY-towers by considering the system obtained by the equivalence relation  $\sim$  on  $\Lambda^i$ ,  $i = 1, 2$  defined as  $x \sim y$  if and only if  $y \in \gamma^s(x)$ . Then on  $\Delta_0^i = \Lambda^i / \sim$  we have the partition  $\mathcal{P}^i = \{\Delta_{0,j}^i\} := \{\Lambda_{0,j}^i / \sim\}$  and the return time function  $R^i : \Delta_{0,j}^i \rightarrow \mathbb{Z}^+$ , and the quadruples  $(F_i, R^i, \mathcal{P}^i, s_i)$ ,  $i = 1, 2$  satisfy conditions (G1)-(G5). Moreover there is natural projection  $\bar{\pi}_i : \mathcal{T} \rightarrow \Delta$  that sends each stable manifold to a point. We can then define the direct product of these two “quotient” GMY-towers and, from previous construction we obtain a new GMY tower for this product. Thus, on  $\Delta_0^1 \times \Delta_0^2$  we have a partition  $\hat{\mathcal{P}}$ , and return time  $T : \Delta_0 \rightarrow \mathbb{N}$  such that for any  $A \in \hat{\mathcal{P}}$  we have  $(F_1 \times F_2)^T(A) = \Delta_0^1 \times \Delta_0^2$ . On the other hand we know that each  $A \in \hat{\mathcal{P}}$  is of form  $A_1 \times A_2$  and  $\bar{\pi}_i^{-1}(A_i) \subset \Lambda^i$ ,  $i = 1, 2$ . Then  $\mathcal{Q} = \{\bar{\pi}_1^{-1}(A_1) \times \bar{\pi}_2^{-1}(A_2) | A_i \in \hat{\mathcal{P}}_i, i = 1, 2\}$  gives the desired partition of  $\Lambda^1 \times \Lambda^2$ . Indeed,  $(f_1 \times f_2)(\bar{\pi}_1^{-1}(A_1) \times \bar{\pi}_2^{-1}(A_2))$  is a  $u$ -subset of  $\Lambda^1 \times \Lambda^2$  because at return times we have  $f_i^T = F_i^T$ ,  $i = 1, 2$ .

All that is left is to check the properties (A1)-(A7). Those that refer to the combinatorial structure follow immediately from the discussion above, others follow immediately from the corresponding properties of the quotient tower. The only new property to check here is the second item in (A3). This follows easily by noticing that from the definition of  $T$  we have  $s_T(x, y) \leq \min\{s_{R^1}(x_1, y_1), s_{R^2}(x_2, y_2)\}$ , where  $s_T, s_{R^1}, s_{R^2}$  denote separation times with respect to return times  $T, R^1, R^2$  and  $x = (x_1, x_2)$ ,  $y = (y_1, y_2)$  and therefore, by the definition of product metric, we have

$$\begin{aligned} \text{dist}_M((f^T)^n(x), (f^T)^n(y)) &= \max_{i=1,2} \{\text{dist}_i((f_i^T)^n(x_i), (f_i^T)^n(y_i))\} \\ &\leq C \beta^{\min\{s_{R^1}(x_1, y_1), s_{R^2}(x_2, y_2)\}} \leq C \beta^{s_T(x, y)}. \end{aligned}$$

Tail estimates transfer directly, since we need to estimate  $m_\gamma\{T > n\}$  for  $\gamma \in \Gamma^u$ .

#### APPENDIX A. TOWER ESTIMATES

We collect here a few simple estimates which hold in general for any Young tower, and which are used mainly in Section 4. For simplicity we state them all for the Tower map  $F$ , the same estimates hold also for  $F'$ . Recall the definition of the partitions  $\eta_n$  in (3.1) and for any  $n \geq 1$ , let

$$\eta_n^0 = \{A \in \eta_n | F^n(A) = \Delta_0\}.$$

**Lemma A.1.** For any  $A \in \eta_n^0$  and  $x, y \in A$  the following inequality holds

$$\left| \frac{JF^n(x)}{JF^n(y)} - 1 \right| \leq D,$$

where  $D$  as in (2.1).



*Proof.* The collection  $\eta_n^0$  is a partition of  $F^{-n}\Delta_0$  and for any  $x \in \Delta_0$  each  $A \in \eta_n^0$  contains a single element of  $\{F^{-n}x\}$ . For  $x \in A$  let  $j(x)$  be the number of visits of its orbit to  $\Delta_0$  up to time  $n$ . Since the images of  $A$  before time  $n$  will remain in an element of  $\eta$ , all the points in  $A$  have the same combinatorics up to time  $n$  and so  $j(x)$  is constant on  $A$ . Therefore  $JF^n(x) = (JF^R)^j(\tilde{x})$ , for the projection  $\tilde{x}$  of  $x$  into  $\Delta_0$  (i.e. if  $x = (z, \ell)$  then  $\tilde{x} = (z, 0)$ ). Thus for any  $x, y \in \Delta_0$  from (2.1) we obtain

$$(A.1) \quad \left| \frac{JF^n(x)}{JF^n(y)} - 1 \right| = \left| \frac{(JF^R)^j(\tilde{x})}{(JF^R)^j(\tilde{y})} - 1 \right| \leq D.$$

□

**Corollary A.2.** For any  $A \in \eta_n^0$  and  $y \in A$  we have

$$(A.2) \quad JF^n(y) \geq \frac{m(\Delta_0)}{m(A)(1+D)}.$$

*Proof.* Lemma A.1 implies  $JF^n(x) \leq (1+D)JF^n(y)$ . Integrating both sides of this inequality with respect to  $x$  over  $A$  gives

$$m(\Delta_0) = \int_A JF^n(x) dm \leq JF^n(y)(1+D)m(A).$$

Hence, for any  $y \in A$  we have the statement. □

**Lemma A.3.** There exists  $M_0 \geq 1$  such that for any  $n \in \mathbb{N}$

$$\frac{dF_*^n m}{dm} \leq M_0.$$

*Proof.* Let  $\nu_n = F_*^n m$ . We will estimate the density  $dF_*^n m/dm$  at different point  $x \in \Delta$  and consider three different cases according to the position of  $x$ . First of all, for any  $x \in \Delta_0$ , from Corollary A.2 we have

$$\begin{aligned} \frac{d\nu_n}{dm}(x) &= \sum_{y \in F^{-n}x} \frac{1}{JF^n y} \leq (D+1) \sum_{A \in \eta_0^{(n)}} \frac{m(A)}{m(\Delta_0)} \\ &\leq (D+1) \frac{m(\Delta)}{m(\Delta_0)} := M_0. \end{aligned}$$

This proves the case  $x \in \Delta_0$ . For  $x \in \Delta_\ell$  with  $\ell \geq n$  we have  $F^{-n}(x) = y \in \Delta_{\ell-n}$ . Since  $JF(y) = 1$  for any  $y \in \Delta \setminus \Delta_0$ ,

$$\frac{d\nu_n}{dm}(x) = \frac{1}{JF^n(y)} = 1.$$

Finally, let  $x \in \Delta_\ell$ ,  $\ell < n$ . Then for any  $y \in F^{-n}x$  the equality  $F^{n-\ell}y = F^{-\ell}x \in \Delta_0$  holds. Hence,  $JF(F^j y) = 1$  for all  $j = n - \ell, \dots, n - 1$ . Therefore by the chain rule we

obtain  $JF^n(y) = JF^{n-\ell}(y)$ . We can reduce the problem to the first case observing

$$\frac{d\nu_n}{dm}(x) = \sum_{y \in F^{-n}x} \frac{1}{JF^n y} = \sum_{y \in F^{\ell-n}(F^{-\ell}x)} \frac{1}{JF^{n-\ell}y} = \frac{d\nu_{n-\ell}}{dm}(F^{-\ell}(x)).$$

This finishes the proof.  $\square$

**Lemma A.4.** For  $n > 0$ , let  $A \in \eta_n^0$  and  $\nu = F_*^n(m|A)$ . Then

$$\left| \frac{\frac{d\nu(x)}{dm}}{\frac{d\nu(y)}{dm}} - 1 \right| \leq D,$$

for any  $x, y \in \Delta_0$ .

*Proof.* By the assumption,  $F^n : A \rightarrow \Delta_0$  is invertible. So for any  $x \in \Delta_0$  there is a unique  $x_0 \in A$  such that  $F^n(x_0) = x$  and  $\frac{d\nu}{dm}(x) = \frac{1}{m(A)} \frac{1}{JF^n x_0}$ . Let  $\varphi = \frac{d\nu}{dm}$  then for  $x, y \in \Delta_0$ , using Lemma A.1 we obtain

$$\left| \frac{\varphi(x)}{\varphi(y)} - 1 \right| = \left| \frac{JF^n(y_0)}{JF^n(x_0)} - 1 \right| \leq D.$$

$\square$

The proof of the following corollary is analogous to the proof of Corollary A.2.

**Corollary A.5.** For  $n > 0$ , let  $A \in \eta_n^0$  and  $\nu = F_*^n(m|A)$ . Then

$$\frac{1}{(1+D)m(\Delta_0)} \leq \frac{d\nu}{dm}(x) \leq \frac{1+D}{m(\Delta_0)}.$$

## APPENDIX B. TAIL ESTIMATES FOR STRETCHED EXPONENTIAL DECAY

Here we prove Lemma 5.8. First we prove the following

**Lemma B.1.** Let  $\tau > 0$ ,  $\theta \in (0, 1)$ . For all  $n \geq (\frac{2}{\tau\theta})^{1/\theta}$  we have

$$\sum_{k \geq n} e^{-\tau k^\theta} \leq \frac{2}{\tau\theta} e^{-\tau n^\theta} n^{1-\theta}.$$

*Proof.* First of all note that we have

$$\sum_{k \geq n} e^{-\tau k^\theta} \leq \int_n^\infty e^{-\tau x^\theta} dx.$$

After change of variables  $t = \tau x^\theta$  we obtain

$$(B.1) \quad \int_n^\infty e^{-\tau x^\theta} dx = \frac{1}{\theta \tau^{1/\theta}} \int_{\tau n^\theta}^\infty e^{-t} t^{1/\theta-1} dt.$$

In [38] it was proved that for any  $a, B > 0$  and  $x > \frac{B(a-1)}{B-1}$

$$(B.2) \quad \int_x^\infty t^{a-1} e^{-t} dt < B x^{a-1} e^{-x}.$$

Substituting (B.2) with  $a = 1/\theta$  and  $B = 2$  into the right hand side of (B.1) finishes the proof.  $\square$

Now, we are ready to prove Lemma 5.8. By Lemma B.1 and definition of  $\bar{\mathcal{M}}_n$ , for sufficiently large  $n$  we have

$$\bar{\mathcal{M}}_{k_j - n_0} \leq C' \sum_{\kappa \geq k_j - n_0} e^{-\tau \kappa^\theta} \leq \frac{2C'}{\tau\theta} e^{-\tau(k_j - n_0)^\theta} (k_j - n_0)^{1-\theta}$$

for any  $0 \leq j \leq i - 1$  and

$$\bar{\mathcal{M}}_{n - n_0 - \sum k_j} \leq \frac{2C'}{\tau\theta} e^{-\tau(n - n_0 - \sum k_j)^\theta} (n - \sum k_j - n_0)^{1-\theta}.$$

Using  $a^\alpha + b^\alpha \geq (a + b)^\alpha$  for  $\alpha \in (0, 1)$  and  $a, b \geq 0$  we obtain

$$\begin{aligned} \bar{\mathcal{M}}_{n - n_0 - \sum k_j} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j - n_0} &\leq \left( \frac{2C'}{\tau\theta} \right)^i e^{\tau(n - in_0)^\theta} (n - \sum k_j - n_0)^{1-\theta} \prod (k_j - n_0)^{1-\theta} \\ &\leq \left( \frac{2C'}{\tau\theta} \right)^i e^{\tau(n - in_0)^\theta} n^{i(1-\theta)}. \end{aligned}$$

Taking  $C_1 = 2C'/\tau\theta$  we obtain the statement in the Lemma.

#### APPENDIX C. DECAY OF CORRELATIONS FOR PRODUCT MEASURES

Our main result can be used in conjunction with [2, Theorem A] to show that rates of Decay of Correlations for a product system follow from rates of Decay of Correlations for the individual components, even without assuming a priori the existence of a Young tower, at least in the polynomial and stretched exponential cases. However, there also exists a much more general, elementary and direct proof of this statement which was told to us by C. Liverani. The calculation is not completely trivial and is clearly of some interest and so, since it does not appear to be in the literature, we reproduce it here with his kind permission.

**Definition 1.** Let  $\mathcal{B}_1, \mathcal{B}_2$  be Banach spaces of measurable observables defined on  $M$ . We denote the *correlation* of non-zero observables  $\varphi \in \mathcal{B}_1$  and  $\psi \in \mathcal{B}_2$  with respect to  $\mu$  by

$$\text{Cor}_\mu(\varphi, \psi \circ f^n) := \frac{1}{\|\varphi\|_{\mathcal{B}_1} \|\psi\|_{\mathcal{B}_2}} \left| \int \varphi(\psi \circ f^n) d\mu - \int \varphi d\mu \int \psi d\mu \right|.$$

We say that  $(f, \mu)$  has decay of correlations at rate  $\{\gamma_n\}$  with respect to  $\mu$  for observables in  $\mathcal{B}_1$  against observables in  $\mathcal{B}_2$  if there exists constant  $C > 0$  such that for any  $\varphi \in \mathcal{B}_1$ ,  $\psi \in \mathcal{B}_2$  the inequality

$$\text{Cor}_\mu(\varphi, \psi \circ f^n) \leq C\gamma_n$$

holds for all  $n \in \mathbb{N}$ .

Rates of decay of correlations have been extensively studied and are well known for many classes of systems and various families of observables, indeed the literature is much too vast to give complete citations. We just mention that the first results on rates of decay of correlations go back at least to [14, 28, 42, 43] in the 70's and since then results have been obtained in [4, 10, 14, 15, 16, 20, 23, 24, 27, 26, 29, 30, 33, 39, 41, 46, 47, 49, 50] amongst others.

Let  $f_i : (M_i, \mu_i) \rightarrow (M_i, \mu_i)$ ,  $i = 1, \dots, p$ , be a family of maps defined on compact metric spaces  $M_i$ , preserving Borel probability measures  $\mu_i$  and let  $f := f_1 \times \dots \times f_p : M \rightarrow M$  be the direct product on  $M := M_1 \times \dots \times M_p$ , and  $\mu := \mu_1 \times \dots \times \mu_p$  be the product measure. We let  $\mathcal{A}_i, \mathcal{B}_i$  denote Banach spaces of functions on  $M_i$  and  $\mathcal{A}, \mathcal{B}$  denote Banach spaces of functions on  $M$ . For completeness we state the minimal requirements on the spaces  $\mathcal{A}_i, \mathcal{B}_i, \mathcal{A}, \mathcal{B}$  needed for the calculations to work. We assume these spaces satisfy the following properties:

- (1) For all  $1 \leq i \leq p$   $\mathcal{A}_i, \mathcal{B}_i \subset L^2(M_i, \mu_i)$  and  $\mathcal{A}, \mathcal{B} \subset L^2(M, \mu)$ .
- (2) For any  $\varphi \in \mathcal{A}, \psi \in \mathcal{B}$  for all  $1 \leq i \leq p$  and  $\mu_1 \times \dots \times \mu_{i-1} \times \mu_{i+1} \times \dots \times \mu_p$  - almost every  $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$  we have  $\hat{\varphi}_i := \varphi(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_p) \in \mathcal{A}_i$ , and  $\hat{\psi}_i := \psi(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_p) \in \mathcal{B}_i$ .
- (3) There exists  $C > 0$  such that for any  $\varphi \in \mathcal{A}, \psi \in \mathcal{B}$  and any  $1 \leq i \leq p$ , we have

$$\|\hat{\varphi}_i\|_{\mathcal{A}_i} \leq C\|\varphi\|_{\mathcal{A}} \quad \text{and} \quad \|\hat{\psi}_i\|_{\mathcal{B}_i} \leq C\|\psi\|_{\mathcal{B}}.$$

It is easy to check that all these conditions are satisfied for some of the commonly considered classes of observables such as Hölder continuous or essentially bounded functions. We now suppose that

$$\text{Cor}_\mu(\varphi_i, \psi_i \circ f_i^n) \leq C\gamma_n$$

for all non-zero  $\varphi_i \in \mathcal{A}_i$  and  $\psi_i \in \mathcal{B}_i$  for all  $i = 1, \dots, p$  and obtain a bound for the correlation function of the product. For simplicity we give the calculation for  $p = 2$ , the general case follows by successive applications of the argument. Let  $\varphi \in \mathcal{A}$  be such that  $\int \varphi d\mu = 0$  and  $\psi \in \mathcal{B}$ . Moreover, let  $\bar{\varphi}(x_1) = \int \varphi(x_1, y) d\mu_2(y)$ . If we fix the first coordinate, by Fubini's theorem we have

$$\begin{aligned} & \int \varphi(x_1, x_2) \psi(f_1^n(x_1), f_2^n(x_2)) d\mu_1 d\mu_2 \\ (C.1) \quad &= \int \left( \int \psi(f_1^n(x_1), f_2^n(x_2)) [\varphi(x_1, x_2) - \bar{\varphi}(x_1)] d\mu_2 \right) d\mu_1 \\ &+ \int \bar{\varphi}(x_1) \psi(f_1^n(x_1), f_2^n(x_2)) d\mu_1 d\mu_2. \end{aligned}$$

Since  $\mu_2$  is  $f_2$ -invariant, we can write the first term of the right hand side as

$$\begin{aligned} I &:= \int \psi(f_1^n(x_1), f_2^n(x_2)) \varphi(x_1, x_2) d\mu_2 - \bar{\varphi}(x_1) \int \psi(f_1^n(x_1), f_1^n(x_2)) d\mu_2 \\ &\leq C\|\varphi(x_1, \cdot)\|_{\mathcal{A}_2} \|\psi(x_1, \cdot)\|_{\mathcal{B}_2} \gamma_n. \end{aligned}$$

The inequality above follows since the left hand side gives the correlations of the second component with respect to  $\mu_2$ . From the third assumption we obtain  $I \leq C\|\varphi\|_{\mathcal{A}}\|\psi\|_{\mathcal{B}}$ . Again by the invariance of  $\mu_2$  we can write the second summand of the equation (C.1) as

$$\int \psi(f_1^n(x_1), f_2^n(x_2))\bar{\varphi}(x_1)d\mu_1d\mu_2 = \int \psi(f_1^n(x_1), x_2)\bar{\varphi}(x_1)d\mu_1d\mu_2.$$

Note that  $\int \bar{\varphi}(x_1)d\mu_1 = 0$  by the choice of  $\varphi$ . Then this expression can be written as

$$\begin{aligned} & \int \left( \int \psi(f_1^n(x_1), x_2)\bar{\varphi}(x_1)d\mu_1 \right) d\mu_2 \\ &= \int \left( \int \psi(f_1^n(x_1), x_2)[\bar{\varphi}(x_1)d\mu_1 - \int \bar{\varphi}(x_1)d\mu_1] \right) d\mu_2 \\ &= \int \left( \int \psi(f_1^n(x_1), x_2)\bar{\varphi}(x_1)d\mu_1 - \int \psi(x_1, x_2)d\mu_1 \int \bar{\varphi}(x_1)d\mu_1 \right) d\mu_2 \end{aligned}$$

The expression under the integral with respect to  $\mu_2$  is exactly the correlation with respect to  $\mu_1$ . Again using the third property of the Banach spaces  $\mathcal{A}$  and  $\mathcal{B}$  we have

$$\text{Cor}_\mu(\varphi, \psi \circ f_1 \times f_2) \leq 2C\gamma_n.$$

This completes the proof of the required statement.

## REFERENCES

- [1] J.F. Alves, C. Dias, S. Luzzatto, Geometry of expanding absolutely continuous invariant measures and the liftability problem, *Ann. Inst. Henri Poincaré, Analyse Non Linéaire*, 30, 101-120, 2013.
- [2] J.F. Alves, J.M. Freitas, S. Luzzatto, S. Vaienti, From rates of mixing to recurrence times via large deviations, *Advances in Mathematics*, 228, 1203-1236, 2011.
- [3] J.F. Alves, S. Luzzatto, V. Pinheiro, Markov structures and decay of correlations for non-uniformly expanding dynamical systems, *Ann. Inst. Henri Poincaré, Analyse Non Linéaire* 22, 817-839, 2005.
- [4] J.F. Alves, D. Schnellmann, Ergodic properties of Viana-like maps with singularities in the base dynamics, *Proceedings of the AMS*, Vol. 141, No. 11, 3943-3955, 2013.
- [5] J.F. Alves, V. Pinheiro, Slow rates of mixing for dynamical systems with hyperbolic structures, *J. Stat. Phys.*, 131, 505-534, 2008.
- [6] J. F. Alves, X. Li, Gibbs-Markov-Young structure with (stretched) exponential recurrence times for partially hyperbolic attractors, Preprint.
- [7] J.F. Alves, D. Schnellmann, Ergodic properties of Viana-like maps with singularities in the base dynamics. *Proc. AMS.* 141, 3493–3955, 2013
- [8] A. Avez, Propriétés ergodiques des endomorphismes dilatants des variétés compactes, *C.R. Acad. Sci. Paris Sér. A-B* 266, A610-A612, 1968.
- [9] V. Baladi, Positive transfer operators and decay of correlations, World Scientific, 2000.
- [10] V. Baladi, S. Gouëzel, Stretched exponential bounds for the correlations of the Viana-Alves skew products. *Second Workshop on Dynamics and Randomness*, Universidad de Chile, 2002.
- [11] M. Benedicks, L. Carleson, The dynamics of the Hénon map. *Ann. Math.*, 122:1-25, 1985.
- [12] M. Benedicks, L.-S. Young, Sinai-Bowen-Ruelle measures for certain Hénon maps, *Invent. Math.*, 112, (1993), 541-576.
- [13] V.I. Bogachev, Measure theory v. 1. Springer, 2006, 500 p.

- [14] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms, Springer Lecture Notes in Math., 75 p., 1975.
- [15] H. Bruin, S. Luzzatto, S. van Strien, Decay of correlations in one-dimensional dynamics, *Annales de l'ENS*, 36, 621-646, 2003.
- [16] J. Buzzi, V. Maume-Deschamps, Decay of correlations on towers with non-Hölder Jacobian and non-exponential return time, *Discrete and Continuous Dynamical Systems*, 12, no. 4, 639-656, 2005.
- [17] J. Buzzi, O. Sester, M. Tsujii, Weakly expanding skew-products of quadratic maps, *Ergod. Th. Dynam. Syst.* 23, 1401-1414, 2003.
- [18] N. Chernov, Statistical properties of piecewise smooth hyperbolic systems in high dimensions. *Disc. and Cont. Dyn. Syst.* 5, 425-448, 1999
- [19] N. Chernov, Decay of correlations and dispersing billiards. *J. Stat. Phys.* 94, 513-556, 1999.
- [20] N. Chernov, R. Markarian, Chaotic billiards. Volume 127 of *Mathematical Surveys and Monographs*. Providence, RI: Amer. Math. Soc., 2006.
- [21] K. Diaz-Ordaz, Decay of correlations for non-Hölder observables for one-dimensional expanding Lorenz-Like maps, *Discrete and continuous dynamical systems*, Vol. 15, no. 1, 159-176, 2006.
- [22] K. Diaz-Ordaz, M.P. Holland, S. Luzzatto, Statistical properties of one-dimensional maps with critical points and singularities, *Stochastics and Dynamics*, Vol. 6, No. 4, 423-458, 2006.
- [23] S. Gouëzel, Decay of correlations for nonuniformly expanding systems, *Bull. Soc. Math. France*, Vol. 134, no. 1, 1-31, 2006.
- [24] F. Hofbauer, G. Keller Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.* 180:119-140, 1982.
- [25] M. Holland Slowly mixing systems and intermittency maps, *Ergodic theory and Dynamical Systems*, Vol. 25, no. 1, 133-159, 2004.
- [26] H. Hu, Decay of correlations for piecewise smooth maps with indifferent fixed points, *Ergodic theory and Dynamical Systems*, Vol 24, no. 2, 495-524, 2004.
- [27] G. Keller, T. Nowicki, Spectral theory, zeta functions and the distributions of points for Collet-Eckman maps, *Comm. Math. Phys.* 149, 31-69, 1992.
- [28] A. Lasota, J.A. Yorke, On the existence of invariant measures for piecewise monotonic transformations, *Trans. of the AMS*, Vol. 186, 1973.
- [29] C. Liverani, Decay of correlations. *Annals Math.* 142 (2), 239-301, 1995.
- [30] C. Liverani, Multidimensional expanding maps with singularities: a pedestrian approach *Ergod. Th. Dynam. Syst.* 33, 168-182, 2013.
- [31] A. Lopes, Entropy and large deviations, *Nonlinearity* 3, 527-546, 1990.
- [32] S. Luzzatto, Stochastic-like behaviour in non-uniformly expanding maps, in *Handbook of Dynamical Systems*, Vol 1B. 265-326, Elsevier, 2006.
- [33] S. Luzzatto, I. Melbourne, Statistical properties and decay of correlations for interval maps with critical points and singularities, *Commun. Math. Phys.* 320, 21-35, 2013.
- [34] V. Lynch, Non-uniformly expanding dynamical systems and decay of correlations for non-Hölder continuous observables, PhD Thesis, University of Warwick, 2003.
- [35] V. Lynch, Decay of correlations for non-Hölder observables, *Discrete and Continuous Dynamical Systems*, 16, no. 1, 19-46, 2006.
- [36] I. Melbourne, M. Nicol, Large deviations for nonuniformly hyperbolic systems, *Trans. Amer. Math. Soc.* 360, 6661-6676, 2008.
- [37] I. Melbourne, M. Nicol, Almost sure invariance principle for nonuniformly hyperbolic systems, *Commun. Math. Phys.*, 260, 131-146, 2005.
- [38] P. Natalini, B. Palumbo, Inequalities for the Incomplete Gamma function, *Mathematical Inequalities & Applications*, 3 (1), 69-77, 2000.

- [39] T. Nowicki, S. van Strien, Absolutely continuous invariant measures for  $C^2$  unimodal maps satisfying the Collet-Eckmann conditions, *Invent. Math.* 93, 619-635, 1988.
- [40] V. Pinheiro, Expanding Measures *Ann. I. H. Poincaré – AN* 28, 889-939, 2011.
- [41] M. Pollicott, M. Yuri, Statistical properties of maps with indifferent periodic points, *Commun. Math. Phys.* 217, 503-520, 2001.
- [42] D. Ruelle, A measure associated with Axiom A attractors. *Amer. J. Math.* 98:619-654, 1976.
- [43] Y. Sinai, Gibbs measures in ergodic theory, *Russ. Math. Surveys*, 27(4) 1972.
- [44] Y. Sinai, Dynamical systems with elastic reflections, Ergodic properties of dispersing billiards. *Uspekhi. Mat. Nauk* 25, 141-192, 1970.
- [45] T. Tao, V.H. Vu Additive combinatorics, *Cambridge studies in advanced mathematics*, v. 105, 2006.
- [46] D. Thomine, A spectral gap for transfer operators of piecewise expanding maps, *Discrete and continuous time Dynamical systems*, Vol. 30, n. 3, 917-944, 2011.
- [47] L.-S. Young, Decay of correlations for certain quadratic maps, *Comm. Math. Phys.* Vol. 146, No. 1, 123-138, 1992.
- [48] L.-S. Young, What are SRB measures, and which dynamical systems have them? *J. Stat. Phys.* Vol. 108, no. 5-6, 733-754, 2002.
- [49] L.-S. Young, Statistical properties of dynamical systems with some hyperbolicity. *Ann. Math.* 147, 585-650, 1998.
- [50] L.-S. Young, Recurrence times and rates of mixing, *Israel J. Math.* 110, 153-188, 1999.